

ISSN 0024-3922

# LINGUISTICA

## LI

LES FRONTIÈRES INTERNES  
ET EXTERNES  
DE LA MORPHOLOGIE

Ljubljana, 2011

Revijo sta ustanovila †Stanko Škerlj in †Milan Grošelj  
Revue fondée par †Stanko Škerlj et †Milan Grošelj

Glavna in odgovorna urednica – Rédactrice en chef  
Martina Ožbot

Številko LI uredil – Responsable du numéro LI  
Gregor Perko

Uredniški odbor – Comité de rédaction  
Janez Orešnik, Stojan Bračič, Gregor Perko

Svetovalni odbor – Comité consultatif  
Wolfgang U. Dressler (Wien)  
Martin Maiden (Oxford)  
Rosanna Sornicola (Napoli)  
Pierre Swiggers (Leuven)

Natis letnika je omogočila  
JAVNA AGENCIJA ZA KNJIGO

Sous les auspices de  
l'AGENCE NATIONALE DU LIVRE

© Univerza v Ljubljani, Filozofska fakulteta, 2011. Vse pravice pridržane.  
© Université de Ljubljana, Faculté des Lettres, 2011. Tous droits réservés.

CIP - Kataložni zapis o publikaciji  
Narodna in univerzitetna knjižnica, Ljubljana

81'1(082)

Les FRONTIERES internes et externes de la morphologie /  
[uredil Gregor Perko]. - Ljubljana : Znanstvena založba Filozofske  
fakultete, 2011. - (Linguistica, ISSN 0024-3922 ; 51)

ISBN 978-961-237-478-5

I. Perko, Gregor

259568896

## PRÉSENTATION

Au cours des dernières décennies où l'on assiste à une refondation de la morphologie, la discipline s'intéresse de plus en plus à des phénomènes « périphériques », « marginaux », « irréguliers » ou « extragrammaticaux », à des phénomènes n'appartenant pas aux « régions nodales » de la morphologie. Le présent numéro de *Linguistica*, premier numéro thématique de notre revue, réunit des études qui s'intéressent aux **frontières internes et externes de la morphologie**. La diversité et la richesse des thèmes abordés et des approches proposées témoignent d'un intérêt croissant que les linguistes, non pas uniquement les morphologues, portent à cette thématique.

Les articles proposés s'articulent autour de cinq axes majeurs.

Un premier axe rassemble les articles qui étudient les frontières entre différentes composantes de la morphologie. L'étude de Michel Roché examine les contraintes lexicales et morphophonologiques sur le paradigme des dérivés en *-aie*. Les résultats de son analyse remettent en question la notion de « règle de construction des lexèmes ». Irena Stramljič Breznik et Ines Voršič se penchent sur les néologismes de sport en slovène et essaient d'évaluer la productivité ou la créativité des procédés morphologiques en jeu. Dans une étude d'inspiration cognitiviste, Alexandra Bagasheva aborde l'hétérogénéité des verbes composés en anglais. En s'appuyant sur les données fournies par les paradigmes verbaux du maltais, Maris Camilleri examine la complexité du phénomène de classes flexionnelles basées sur les radicaux. Trois travaux concernent des procédés typiquement « extragrammaticaux » servant à former le plus souvent des occasionnalismes : Arnaud Léturgie tente de dégager des propriétés prototypiques de l'amalgamation lexicale en français, notamment celles de la création des mots-valises ; Silvia Cacchiani analyse, à l'intérieur du cadre de la morphologie naturelle, les mots-valises formés à partir de noms propres et de substantifs, phénomène relativement récent en italien ; Thomas Schwaiger, se fondant sur les données tirées de la base « Graz Database on Reduplication », traite des constructions rédupliquées dans une perspective universelle. Deux articles adoptent une perspective contrastive : celui d'Eva Sicherl et Andreja Žele étudie la productivité des diminutifs nominaux en slovène et en anglais ; celui de Marie-Anne Berron et Marie Mouton propose une analyse détaillée de l'importance quantitative et qualitative des procédés de la morphologie marginale dans le slam en France et en Allemagne. Géraldine Walther présente un nouveau modèle général, inscrit dans une approche réalisationnelle, qui permet d'évaluer et de formaliser la (non-)canonicité de phénomènes flexionnels.

Un deuxième axe se situe sur la frontière séparant la morphologie dérivationnelle de la morphologie flexionnelle. Tatjana Marvin se penche sur le problème de la préservation de l'accent dans les dérivés anglais, mettant en comparaison les approches de Chomsky et de Marantz avec celle de la théorie de l'Optimalité. L'article de Patrizia Cordin explore la façon dont les constructions locatives verbales dans les langues et les dialectes romans, notamment dans le dialecte trentin, ont perdu en partie leur sens spatial au profit d'un sens grammatical plus abstrait (aspect, résultat, intensité etc.). Se

focalisant sur la langue indigène d’Australie kayerdild, Erich Round insiste sur l’utilité de la notion de « morphème » qui permet d’instaurer un niveau de représentation linguistique intermédiaire entre les niveaux lexical, morphosyntaxique et morphophonologique. La contribution de Varja Cvetko Orešnik, qui adopte le cadre théorique de l’école ljubljanaise de la syntaxe naturelle, apporte quelques données nouvelles sur la morphophonologie et la morphosyntaxe du verbe en ancien indien.

Les deux articles suivants examinent des questions d’ordre morphologique auxquelles se heurtent les langues entrant en contact : Georgia Zellou traite du cas du circumfixe /ta...-t/ que l’arabe marocain a emprunté au berbère, tandis que Chikako Shigemori Bučar analyse le sort réservé aux emprunts japonais en slovène.

L’axe diachronique est exploré principalement par les articles de Douglas Lightfoot, de Javier E. Díaz Vera et de Metka Furlan. D. Lightfoot examine la pertinence de la notion d’affixoïde et évalue la « suffixoïdité » de l’élément germanique « -mann ». J. E. Díaz Vera décrit la lexicalisation, voire la grammaticalisation des verbes causatifs en ancien anglais. Metka Furlan se penche sur une relation « morphologique » archaïque rattachant le nom protoslave *pol’e* à l’adjectif hittite *palhi-*.

Autour du dernier axe se réunissent les contributions qui examinent la frontière entre la morphologie et les disciplines linguistiques voisines, telles que la phonologie ou la syntaxe. Marc Plénat étudie en détail les contraintes morphologiques, syntaxiques et phonologiques pesant sur la liaison de l’adjectif au masculin singulier sur le nom en français. Janez Orešnik, fondateur de l’école ljubljanaise de la syntaxe naturelle, explore le comportement morphologique de l’impératif dans une perspective universelle. Mojca Schlamberger Brezar propose une étude contrastive, à partir des données tirées de corpus monolingues et parallèles, de la grammaticalisation du gérondif et du participe en français et en slovène. L’article de Gašper Ilc applique la notion de « cycle de Jespersen » à la négation dans le slovène standard et dans les dialectes pannoniens. Mojca Smolej propose une étude fouillée de l’émergence des articles défini et indéfini dans le slovène parlé spontané.

Gregor Perko  
Université de Ljubljana

## PRESSION LEXICALE ET CONTRAINTES PHONOLOGIQUES DANS LA DÉRIVATION EN *-AIE* DU FRANÇAIS

La dérivation qui donne les *chênaie*, *roseraie*, *châtaigneraie* ne semble pas poser de problèmes particuliers. Remarquablement homogène sur le plan sémantique, elle construit au moyen du suffixe *-aie* des noms de collectifs (plantations, milieux naturels) à partir de noms de plantes – des arbres dans la majorité des cas. Quelques formes, cependant, ne laissent pas d'intriguer : *cerisaie*, *amandaie*, qui semblent construits sur le nom du fruit plutôt que sur celui de l'arbre ; *joncheraie*, *ronceraie*, où un segment intermédiaire vient s'intercaler entre la base et le suffixe ; *boulaie*, *prunellaie*, *noiseraie*, dont les radicaux s'écartent de la forme attendue. Le corpus utilisé par Tomassone et Combettes (1970) et par Diekmann (1975) pour leurs comparaisons avec la dérivation en *-ière* – une quarantaine de formes – permettait de formuler des hypothèses pour expliquer ces écarts, guère de les vérifier. En exploitant les ressources de la Toile, on peut aujourd'hui rassembler plus de 400 formes bien attestées, construites sur quelque 270 bases différentes.<sup>1</sup> Le gain de un à dix enrichit considérablement la vision qu'on peut avoir de cette dérivation. Elle se diversifie, fait apparaître des difficultés nouvelles, révèle la créativité actuelle et, surtout, en mettant les écarts en perspective, elle permet de les interpréter de façon moins hasardeuse.

On envisagera d'abord les questions touchant à l'allomorphie radicale (choix du thème et contraintes phonologiques conduisant à modifier le radical), puis l'influence des paradigmes lexicaux – série dérivationnelle et famille dérivationnelle – dans lesquels le dérivé est destiné à s'intégrer. Ces observations, parallèlement à d'autres recherches, conduisent à remettre en cause l'approche traditionnelle de la dérivation en termes de « Règles de Construction des Mots ».

### 1. APPROCHE MORPHOPHONOLOGIQUE

Parmi les phénomènes dans lesquels se manifeste l'allomorphe radicale, certains sont intrinsèques au lexème base, qui se présente sous plusieurs formes parmi lesquelles la dérivation doit choisir (§ 1.1). D'autres tiennent à l'adaptation du thème sélectionné en fonction de la dérivation concernée (§ 1.2).

---

\* Adresse de l'auteur : Le Péloy, Estantens, 31600 Muret, France. Mél : mroche@univ-tlse2.fr

<sup>1</sup> La base de données (DiMoC-aie.mdb et DiMoC-aie.xls) est consultable sur <http://w3.erss.univ-tlse2.fr:8080/index.jsp?perso=mroche>. On y trouvera les attestations (contexte, références) des dérivés qui ne figurent pas dans les dictionnaires.

## 1.1. Le choix du thème

Dans une perspective lexématique, on considérera, à la suite d'Aronoff (1994), qu'un lexème est représenté dans le lexique par un ensemble de thèmes qui constituent ce que Bonami et Boyé (2005, entre autres) ont appelé son « espace thématique ». Chaque thème y est alloué à une ou plusieurs cases de ses paradigmes flexionnel et dérivationnel. Pour l'adjectif ROND, par exemple, les thèmes fondamentaux sont un thème A /rɔ̃/ sélectionné par la forme libre masculine (*rond*) et un thème B /rɔ̃d/ sélectionné par la forme libre féminine (*ronde*) et pour la dérivation (*rondeur*). Le dérivé *rotondité* révèle en plus un thème supplétif savant /rɔ̃d/ emprunté au latin (nous l'appellerons S), logiquement sélectionné par le suffixe « savant » *-ité* alors que le suffixe « populaire » *-eur* sélectionne le thème B. Pour certains lexèmes, on trouve en outre un thème supplétif populaire qui garde la trace d'apophonies anciennes. NOUVEAU, par exemple, est représenté dans le radical de *renouveler* non par le thème B /nuvəl/ mais par un thème C /nuvəl<sup>2</sup>/.

Le suffixe *-aie* est un suffixe « populaire ». Transmis depuis le latin *-eta* (pluriel de *-etum*) par la voie populaire, il sélectionne dans tous les dérivés du corpus dictionnaire un thème populaire, en général un thème B identique par défaut au thème A (CHÊNE /ʃɛn/ ~ /ʃɛn/ → *chênaie*) ou révélant une consonne latente (HOUX /u/ ~ /us/ → *houssaie*). Pour quelques bases en *-eau* et les nombreuses bases en *-ier*, cependant, le thème sélectionné est un thème C en /-əl/ (FOUTEAU /futo/ ~ /futəl/ ~ /futəl/ → *foutelaie*) ou en /-ər/ (ROSIER /rɔzje/ ~ /rɔzjer/ ~ /rɔzər/ → *roseaie*) que l'on retrouve devant d'autres suffixes dans les dérivations anciennes (CHAPEAU → *chapelier*, CHEVALIER → *chevalerie*).

Particularité remarquable : alors que les allomorphies qui donnent les thèmes C sont depuis longtemps caduques, les dérivés en *-aie* les plus récents construits sur une base en *-ier* ont toujours une finale en *-eraie*. Sur CHAPEAU, à côté de *chapelier*, le plus récent *chapeauter* et la *tresse chapeautière* des *Exercices de style* de Queneau utilisent non le thème C mais un thème B identique au thème A (complété par un /t/ épenthétique). Sur des bases en *-ier*, si CHEVALIER donne *chevaleresque*, GAUTIER, CLAVIER, POMPIER, VANIER donnent *gautièresque* (et non *°gauteresque*), *clavièresque*, *pompièresque*, *vanièresque*. ARGANIER, KAPOKIER, PALÉTUVIER, RONIER, en revanche, donnent *arganeraie* (et non *°arganiéraie*), *kapokeraie*, *palétuveraie*, *roneraie*. Et ainsi de plusieurs dizaines d'autres qui, très vraisemblablement, ne remontent pas au-delà du 19<sup>e</sup> siècle (absents, pour la plupart, des dictionnaires, il est difficile de les dater). Dans la base de données de dérivés en *-esque* réunie par Marc Plénat<sup>3</sup>, on trouve quatorze exemples du type POMPIER → *pompièresque* et les seules exceptions<sup>4</sup> sont elles-

<sup>2</sup> Pour plus de précisions sur l'espace thématique des nominaux, voir Roché (2010).

<sup>3</sup> Base de données annexée à l'ouvrage de B. Habert *Instruments et ressources électroniques pour le français* (Paris, Ophrys, 2005) et consultable sur le site [http://www.toeditions.com/Sources/Habert\\_Bases-de-donnees-2.htm](http://www.toeditions.com/Sources/Habert_Bases-de-donnees-2.htm).

<sup>4</sup> On trouve également un autre traitement (qui ne nous intéresse pas ici) consistant à superposer le suffixe à la finale /e/ de la base : (MADAME) RECAMIER → *recamiesque*.

mêmes révélatrices. Outre *chevaleresque*, il s'agit de (JACQUES) BERGIER → *bergeresque*, CAVALIER → *cavaleresque*, OLIVIER → *oliveresque* et PAPIER → *paperesque*, influencés respectivement par *bergerie*, *cavalerie*, *oliveraie* et *paperasse*. Les allomorphies /je/ ~ /əɾ/ des finales en *-ier(e)*, dans les dérivés récents, ne sont pas le résultat d'une règle phonologique, par conséquent, mais celui d'une pression lexicale qui s'exerce, pour les uns, au sein de la série dérivationnelle des dérivés en *-(er)iaie*, pour les autres au sein des familles dérivationnelles concernées (le patronyme *Bergier* étant pour l'occasion assimilé au nom commun *berger*). Nous y reviendrons.

A côté de ces dérivés construits sur un thème populaire, beaucoup d'autres, totalement absents des dictionnaires généralistes mais bien attestés dans les publications spécialisées, sont construits sur un thème savant. Tous les degrés sont représentés, de l'allomorphie au supplétisme, suivant que le nom latin de la plante reste proche du nom français ou s'en éloigne plus ou moins complètement : *cariçaie* (← lat. *carex* ~ *caric(is)* 'carex'), *euphorbiaie* (← lat. *euphorbi(a)* 'euphorbe'), *rosmarinaie* (← lat. *rosmarin(us)* 'romarin'), *lauraie* (← lat. *laur(us)* 'laurier'), *viburnaie* (← lat. *viburn(um)* 'viorne'), *lariçaie* (← lat. *larix* ~ *laric(is)* 'mélèze'), *sambuçaie* (← lat. *sambuc(us)* 'sureau'), *spartiaie* (← lat. *spartium junceum* 'genêt d'Espagne'), *suberaie* (← lat. (*quercus*) *suber* 'chêne-liège'), *typhaie* (← lat. *typh(a)* 'massette'), etc. Nombre de ces dénominations viennent doubler un dérivé construit sur le thème populaire : *abiétaie* / *sapin(er)iaie*, *acéraie* / *éabl(er)iaie*, *arbutaie* / *arboseaie*, *bétulaie* / *boulaie*, *buxaie* / *buisaie*, *cératoniaie* / *carouberaie*, *cupressaie* / *cypressaie*, *fraxinaie* / *frênaie*, *juniperaie* / *génévraie*, *lavandulaie* / *lavand(er)iaie*, *pistaciaie* / *pistacheraie*, *saliçaie* / *saulaie*, *sorbaie* / *sorberaie*, *taxaie* / *iv(er)iaie*, *tiliaie* / *tillaie*, *ulmaie* / *orm(er)iaie*, etc.

On peut penser que ces dérivés sont construits directement sur le nom latin des plantes, couramment utilisé par les naturalistes. Certains contextes font d'ailleurs apparaître côte à côte ce nom latin et celui du biotope dont la plante est caractéristique :

- (1) Rosmarinaie à *Rosmarinus officinalis*. Junipéraie à *Juniperus phoenicea*.  
[physio-geo.revues.org/696?file=1](http://physio-geo.revues.org/696?file=1)

Il n'en reste pas moins qu'il s'agit de mots français et que dans *bétulaie* comme dans *boulaie* le suffixe est le même. Ce qui tend à remettre en cause le principe selon lequel une seule et même case de l'espace thématique serait sélectionnée par une dérivation donnée. La correspondance entre espace thématique et paradigme flexionnel est rigoureusement univoque, elle l'est beaucoup moins pour la dérivation.

## 1.2. Les contraintes phonologiques

Le thème fourni par l'espace thématique du lexème base n'est pas toujours conservé tel quel, il peut être modifié pour constituer le radical du dérivé (Roché 2010). Ces aménagements – troncations, ajouts, substitutions... – visent, pour une part, à satisfaire des contraintes phonologiques de « bonne formation ».

Lorsque la finale *-ier* de la base est précédée par un /r/ (dans *poirier*, par exemple), l'allomorphie radicale donnerait dans le dérivé une finale *-reraie* (*°poireraie*) dyspho-

nique du fait de la consécution trop rapprochée du même phonème. Une contrainte dissimilative (Plénat 2011) réduit le radical à *poir-* dans le dérivé effectivement réalisé *poiraie*. Même chose dans *mûrier* → *mûraie*, *coudrier* → *coudraie*, *genévrier* → *genevraie*. Ces troncations ne sont pas obligatoires, cependant : on trouve aussi *mûreraie*, *poivre-raie*, *gingembre-raie* (et même *poireraie* et *coudreraie*, dans des toponymes).

Si le /r/ est plus éloigné de la finale, la contrainte dissimilative est moins forte mais elle explique vraisemblablement le choix de *cerisaie* plutôt que *ceriseraie* (qui n'est attesté que très marginalement). Sur *prunier*, à côté de *prunaie* où, comme dans *cerisaie*, le radical est tronqué, *prunelaie* représente une autre solution : le /r/ est remplacé par un phonème voisin mais suffisamment distinct pour éviter la répétition.

Inversement, la finale en /əl/ du thème C de *bouleau*, qui donnerait °*boulelaie* (cf. *fouteau* 'hêtre' → *foutelaie*), est remplacée par /ər/ dans *boulaie*. Elle peut aussi être tronquée pour donner *boulaie*, comme celle en /ər/ dans *prunaie*. On retrouve les mêmes traitements pour *tilleul* → *tillaie* / *tilleraie*, préférés à °*tillelaie* à une époque où le -ll- de *tilleul* était encore la latérale /ʎ/, trop proche du /l/ final.

Lorsque le thème sélectionné, tel qu'il est fourni par l'espace thématique, n'offre pas d'accroche consonantique, c'est la contrainte antihiatique qui conduit à le modifier. À partir d'un thème populaire, la solution la plus fréquente consiste à introduire une consonne épenthétique. Celle-ci peut être un /t/, épenthèse par défaut du français (*caoutchouterai*, *alfaterai*, *bambout(er)ai*, *caféterai*), mais aussi un /z/ (*bambous(er)ai*) et surtout un /r/ (*bambourai*, *acaciaraie*, *hévéraraie*, *thuyaraie*, *kiwiraie*, *caféraie*, *cacaoraie*). On trouve également une épenthèse en /r/ dans *bouleaurai*, attesté marginalement à côté des plus anciens *boulaie* et *boulaie*. Comme dans *chapeauter*, la conservation de la finale y manifeste l'obsolescence des allomorphes -eau ~ -el-. Épenthèse en /r/ également dans *rhodoraie*, où la base *rhododendron* est remplacée par son abréviation *rhodo*.

La solution alternative à l'épenthèse, pour les thèmes à finale vocalique, est la troncation de la voyelle : *mimos(er)ai*, *nymphéai*, *orchideraie*. Voire de plusieurs voyelles (*nymphaie* ← *nymphéa*) ou de semi-consonne + voyelle (*gonakaie* ← *gonakié*<sup>5</sup>). En contexte savant, sur une base latine, la troncation de la voyelle thématique est de règle, comme en latin : *oryza* 'riz sauvage' donne régulièrement *oryzaie*. Ce n'est pas le même *acacia* qui est à la base de *acaciaie* et de *acaciaraie* : le premier est une dénomination latine de la nomenclature, le second un mot français (et il ne s'agit pas du même arbre). Mais ce principe n'est pas toujours respecté : *thuyaie* et *thuyaraie* ont été relevés dans le même type de contexte.

On a pu établir d'autre part (Plénat 2009) qu'une contrainte prosodique favorisait, dans les mots construits, les radicaux ayant au moins deux syllabes par rapport aux radicaux monosyllabiques. Parmi les dérivés enregistrés par les dictionnaires, dont on peut supposer qu'ils sont les plus anciens ou les mieux établis, la plupart des formes en -eraie, en dehors de celles qui ont pour base un nom en -ier, sont construites sur des monosyllabes : *jonc* → *joncheraie*, *ronce* → *ronceraie*, *charme* → *charmeraie*, *pin* →

---

<sup>5</sup> *Gonakié* est un doublet de *gonakier*, dont le dérivé (régulier) est *gonakeraie*.



*pineraie, teck* → *teckeraie*. Il en va de même pour de nombreux toponymes (*L'Aulneraie, La Buisseraie, L'Herberaie, La Housseraie, La Sauleraie, La Vergneraie...*), dont certains au moins doivent être anciens. On peut interpréter le segment *-er-* comme un interfixe (Roché 2003), élément neutre (sémantiquement) qui vient s'intercaler entre le thème tel qu'il est fourni par l'espace thématique de la base et le suffixe. Le même interfixe que dans *nappe* → *napperon, Pâques* → *pâquerette, hache* → *hachereau, fève* → *féverole, sec* → *sécheresse*, etc. où la suffixation est « décalée » (Plénat et Roché 2004) de la même façon. Dans les couples *jonchaie / joncheraie, saulaie / sauleraie*, etc., la première forme satisfait la contrainte de fidélité (rien n'est ajouté à la base ou au suffixe), la seconde satisfait la contrainte prosodique (le radical est dissyllabique).

## 2. APPROCHE LEXICALE

Tout mot construit appartient par définition à deux paradigmes : une *famille dérivationnelle*, pendant synchronique de la « famille de mots » étymologique, qui réunit tous les lexèmes construits directement ou indirectement sur une même base ; une *série dérivationnelle*, qui réunit tous les lexèmes construits selon un même modèle. Il suffit de deux éléments dans l'une (*parentèle* et *parent*, par exemple) et de deux éléments dans l'autre (*parentèle* et *clientèle*) pour que *parentèle* soit un mot construit, si la relation (formelle, sémantique, catégorielle) qui unit *parentèle* à *parent* est identique à celle qui unit *clientèle* à *client*. Et reproductible : sur ce modèle qui semblait en sommeil depuis le latin on a récemment formé *patientèle* pour désigner l'ensemble des patients d'un médecin. Mais il n'est sans doute pas fortuit que le lexème ayant bénéficié de cette résurrection soit *patient*, doté de la même finale que *client* et *parent* : la maigre série dérivationnelle *clientèle, parentèle, patientèle* est d'autant mieux perçue que ses éléments ne partagent pas seulement le rare suffixe *-èle* mais une chaîne plus étendue. Dans la dérivation en *-aie*, plus productive mais relativement circonscrite, la double dimension paradigmatique joue également un rôle déterminant.

### 2.1. L'influence de la série dérivationnelle

Parmi les bases des noms de collectifs en *-aie*, beaucoup sont des noms d'arbres en *-ier*. De ce fait, les finales en *-eraie* sont particulièrement nombreuses parmi les dérivés. Dans la série dérivationnelle, les formes en *-eraie* constituent par conséquent une sous-série attractive : le mécanisme d'analogie qui est au fondement même de la dérivation tend à généraliser cette finale à l'ensemble des dérivés. Le locuteur qui forge *mandarinaie*, ou qui le rencontre pour la première fois, n'ajoute pas le suffixe *-aie* à *mandarinier*. Il produit, ou saisit, *mandarinaie* comme le résultat d'une quatrième proportionnelle où ont pris place des lexèmes déjà présents dans sa mémoire lexicale – *oranger* et *orangeraie* sur un axe, par exemple, *oranger* et *mandarinier* sur l'autre.

(2) <i>oranger</i>	:	<i>orangeraie</i>
<i>mandarinier</i>	:	<b><i>mandarinaie</i></b>

Sur la lancée de cette sous-série, il forgera *bambouseraie*, ou ne sera pas surpris s'il le rencontre, sans prendre garde que le nom du végétal concerné n'est pas *bambousier* mais *bambou*.

La tendance à étoffer la marque caractéristique de la série lexicale est d'autant plus forte que le suffixe *-aie* est phonologiquement un suffixe pauvre, réduit à un seul phonème. Il a d'ailleurs été confondu dans le passé avec son presque homonyme *-ée*, d'où les *fragnée*, *saulée*, *sapinée*, etc. qui sont en fait des *fragnaie* (*frênaie*), *saulaie*, *sapinaie*<sup>6</sup>. L'effet de rime propre à la dérivation suffixale est ramené ici à une simple assonance. D'où le besoin de remplacer cette rime à peine « suffisante » par une rime plus riche qui renforcera la cohérence de la série et la perception du mot construit en tant que tel. Avec un suffixe phonologiquement plus étoffé mais qui entre dans des dérivations beaucoup plus diversifiées, le suffixe *-isme*, on observe de la même façon une tendance à la constitution de sous-séries en *-icisme*, *-acisme*, *-alisme*, *-ianisme*, suivant les domaines (Roché 2011). Tendance qui entraîne des écarts par rapport à ce que donnerait la simple concaténation de *-isme* au thème fourni par la base : *rho* → *rotacisme*, *Flandres* / *flamand* → *flandricisme*, *personne* → *personnalisme*, *Hegel* → *hégélianisme*...

L'obtention de la finale caractéristique d'une sous-série peut se faire de deux manières : par une modification du radical ; par la substitution d'une variante à la forme standard du suffixe.

### 2.1.1. Les modifications du radical

Les modifications du radical que nous avons énumérées plus haut répondaient à des règles phonologiques ou à des contraintes de bonne formation. Mais la plupart ont aussi pour résultat d'aboutir à une finale en *-eraie* (ou au moins en *-raie*). Rappelons-les brièvement :

- Pour les bases en *-ier*, l'allomorphie /je/ ~ /ɛr/ a été conservée jusqu'à aujourd'hui (*ronier* → *roneraie*) alors qu'elle est caduque dans les autres dérivations. On observe parallèlement que dans les dérivations « savantes » les thèmes latins en /ɛr/ (e.g. *acer*, *suber*, *juniper*-) donnent tantôt des formes en *-éraie* (*acéraie*, *subéraie*, *junipéraie*...), tantôt des formes en *-eraie* (*aceraie*, *suberaie*, *juniperaie*...). Dans les premières, l'allomorphie obéit à une règle purement phonologique : l'alternance /ɛ/ ~ /e/, comme dans *cancer* → *cancéreux* ou *prison* / *carcer* → *carcéral* (on ne pourrait pas avoir \**cancereux* ou \**carceral*). Dans les secondes, l'alternance /ɛ/ ~ /ə/ résulte du conditionnement lexical : elle rapproche le dérivé de la série en *-eraie*.
- L'épenthèse privilégiée est la consonne /r/ (*acacia* → *acaciaraie*) alors qu'elle est assez rare d'une façon générale.

---

<sup>6</sup> Sur cette question, voir Alexander (1912 : 28), Nyrop (III : 89-90), Tomassone et Combettes (1970 : 225). Ces dérivés en *-ée* sont surtout attestés en ancien français, mais Nyrop signale que « G. Flaubert écrit *hêtrée* dans le premier chapitre de *Madame Bovary* ».

- Le seul interfixe utilisé est l'interfixe *-er-* (*jonc* → *joncheraie*) alors que devant d'autres suffixes on en trouve généralement plusieurs. Devant *-on*, par exemple, en plus de *-er-* (*napperon*), il y a souvent *-ill-* (*bottillon*), *-ich-* (*pâlichon*), *-et-* (*gueueleton*), etc.

On peut ajouter quelques observations plus ponctuelles :

- la troncation de *rhododendron* dans *rhododendraie*, qui permet d'obtenir une finale en *-raie*<sup>7</sup>;
- le mfr. *milleraie* → *millet*, où la finale *-eraie* se substitue à *-etaie* sans nécessité phonologique ;
- le moindre impact de la contrainte dissimilative sur les bases en *-rier* : le type *poirier* → *poiraie*, *genévrier* → *genevraie* semblait l'emporter sur *poireraie* (beaucoup plus rare) ou °*genèvreraie* tandis que pour des formations plus récentes on trouve plutôt *poivreraie* que *poivraie*, *gingembreraie*, *laureraie* (dans la légende d'une photo représentant une allée de lauriers-roses<sup>8</sup>).

### 2.1.2. La variante *-eraie*

L'apparition et la diffusion de la variante *-eraie* résultent d'un processus bien connu de réanalyse, favorisé par l'allomorphie /je/ ~ /ər/ des noms d'arbres en *-ier* qui altère la reconnaissance de la base. Un dérivé comme *châtaigneraie* est perçu non comme *châtaignier* + *-aie* mais comme *châtaigne* + *-eraie*. C'est la même fausse coupe qui, dans un contexte phonologique semblable, a produit *-erie*, d'abord variante de *-ie* puis forme standard du suffixe. On a pu construire *gendarmérie* sur *gendarme* et *ébénisterie* sur *ébéniste* comme auparavant *mairie* sur *maire* et *chevalerie* sur *chevalier* parce que ce même *chevalerie* avait été perçu comme *cheval* + *-erie*. La réanalyse, cependant, n'implique pas que les locuteurs aient jamais associé *chevalerie* à *cheval* en tant que lexème porteur de sens. Sémantiquement, le court-circuit ne fonctionne pas. Le processus est purement formel et résulte d'un double effet paradigmatique : la récurrence de la finale *-erie* dans la série dérivationnelle, la présence du lexème *cheval* dans la famille dérivationnelle de *chevalier* et *chevalerie*.

La variante *-eraie* a dû apparaître très tôt : l'ancien français *jardinaie* en témoigne. Dans la plupart des formes du corpus dictionnaire où l'on pourrait la reconnaître, la séquence *-eraie* est aussi interprétable comme interfixe + suffixe : *joncheraie*, *ronceraie*, *charmeraie*, *pineraie*, *teckeraie* ont des bases monosyllabiques qui justifient un décalage de la suffixation. Mais ce n'est pas le cas de *bamboueraie*, *cactuseraie*, *concombreaie*, *douglasseraie*, *gentianeraie*, *lavanderaie*, *mélèzeraie*, *mimoseraie*, *plataneraie*, *sisaleraie*, etc. (radicaux dissyllabiques), *artichauteraie*, *baobaberaie*, *champignonneraie*, *eucalypteraie*, *margueriteraie*, *pissenliteraie*, etc. (radicaux trisyllabiques).

<sup>7</sup> Cette troncation est rendue possible par la longueur de la base. Il ne semble pas cependant qu'une contrainte de taille tende à accourcir les bases longues. Seul *micocoulaie* (← *micocoulier*, à côté de *micocouleraie*) pourrait correspondre à une telle tendance.

<sup>8</sup> Le doublet *lauraie* mentionné plus haut est également récent mais il apparaît dans un contexte savant et doit être construit, vraisemblablement, sur le latin *laurus* comme *sorbaie* l'est sur *sorbus* (face au populaire *sorberaie*).

La fréquence des formes en *-eraie* dans les formations récentes (cf. autres exemples *infra* § 3) laisse supposer que, pour certains locuteurs au moins, *-eraie* est devenu la forme standard du suffixe.

### 2.1.3. Le paradigme des noms d'arbres en *-ier*

Une autre série lexicale est à prendre en compte : celle des noms d'arbres (ou d'autres végétaux) en *-ier*. Normalement, ils sont construits sur le nom du fruit (ou d'une autre production) : *pomme* → *pommier*, *rose* → *roseraie*. D'un point de vue constructionnel, c'est une relation actancielle qui associe le dérivé à sa base : le pommier fait des pommes comme le potier fabrique des pots, le pétrolier transporte le pétrole, la vache laitière donne du lait, etc. Le dérivé désigne l'agent ou l'instrument (ou une autre entité assimilée à un agent) d'un procès dont la base représente le patient ou le résultat. Les dérivés peuvent appartenir à de multiples classes référentielles (humains, animaux, instruments, etc.). A cause de la sous-série, ancienne et nombreuse, des noms d'arbres fruitiers, cependant, la finale *-ier* a été perçue comme une marque caractéristique du nom d'arbre en lui-même. D'où l'utilisation du suffixe comme « intégrateur paradigmatique » (Corbin 1987). Dans deux types de situations, principalement :

- (1) Un nom d'arbre hérité du latin est déjà dans la langue, il est doublé par un dérivé synonyme. C'est le type afr. *peuple* → *peuplier*. Dans ce cas, ou bien le dérivé élimine le primitif, après un temps de cohabitation plus ou moins long (*peuplier*, *coudrier*, *églantier*...); ou bien il ne s'impose pas mais subsiste plus ou moins durablement, surtout dans les parlers régionaux (*ormier*, *platanier*, *aubépinier*...). Le processus est toujours à l'œuvre : on trouve des attestations de *teckier*, *acacier*, *mimosier*, *baobabier*... C'est la coexistence des deux synonymes, dans la famille dérivationnelle de la base, qui a une conséquence sur la forme des dérivés en *-aie* (*infra*, 2.2.2).
- (2) Le suffixe *-ier* sert à « habiller » un emprunt. C'est le type *palétuvier* (d'abord *appariturier*, du tupi *aparahiwa*), qui n'a pas d'incidence sur la dérivation en *-aie*.

## 2.2. L'influence des familles dérivationnelles

La dérivation fait entrer le dérivé dans la famille dérivationnelle de sa base. Elle instaure ainsi une relation non seulement entre lui et un lexème donné, sa base proprement dite, mais aussi avec les autres éléments qui constituent sa famille dérivationnelle. En particulier les plus proches : adjectif de relation pour le nom, nom de qualité pour l'adjectif, nom processif pour le verbe. Quelle est la base d'*échangisme* ? Le nom *échange* ou le verbe *échanger* ? On peut, pour des raisons de cohérence catégorielle, préférer l'un à l'autre, mais sémantiquement le choix est indifférent. Quand il n'y a pas homophonie, la forme dirige automatiquement vers un lexème donné : *lepénisation* ne peut avoir pour base que *lepéniser*. Mais le verbe est beaucoup plus rare que son déverbal, qui est perçu comme désignant la diffusion en profondeur des idées de Le Pen, donc directement en relation avec le primitif. La forme, d'autre part, peut être trompeuse : dans les exemples mentionnés plus haut, le *personnalisme*

est une philosophie fondée sur la *personne*, pas sur ce qui est *personnel*, l'*hégélianisme* est la philosophie de *Hegel*, pas celle des *hégéliens*. L'*américanisme* est une prédilection pour l'Amérique aussi bien que pour ce qui est américain (il n'y a pas d'*américisme*), le *japonisme* une prédilection pour ce qui est japonais aussi bien que pour le Japon (il n'y a pas de *japonaisisme*). Pour des raisons à la fois phonologiques et lexicales, les dérivés de cette sous-série sont construits formellement tantôt sur le nom de pays tantôt sur l'adjectif ethnique (Roché 2008), mais ils renvoient systématiquement à la fois au nom de pays, au nom des habitants, au nom de langue quand il y en a un, et aux adjectifs de relation correspondant à chacun de ces trois noms.

Dans la dérivation en *-aie*, quatre cas de figure peuvent se présenter :

### 2.2.1. Le nom de la plante est seul face à son dérivé

Si le nom de collectif est le premier dérivé apparaissant dans la famille, ou si d'autres dérivés antérieurs n'interviennent pas dans la dérivation qui nous intéresse,

(3) <i>rhododendron</i>	:	<i>rhododendraie</i>
<i>lilas</i>	:	<i>lilaserai</i>

un écart par rapport à la forme attendue ne peut avoir qu'une raison morphologique : modification du radical (*rhododendraie* au lieu de *°rhododendronaie*), variante du suffixe (*lilaserai* au lieu de *°lilasaie*).

### 2.2.2. Le nom de la plante se présente sous plusieurs formes

Il s'agit principalement du type *peuple* / *peuplier*, *coudre* / *coudrier* dont nous avons vu plus haut l'origine. Il faut y ajouter les doublets formés avec d'autres suffixes (*orme* / *ormeau*, afr. *boul* / *bouleau*, afr. *fou* / *fouteau* 'hêtre', afr. *sap* / *sapin*), parfois dès le latin (*teil* / *tilleul*, plus l'afr. *tillier*).

Chacune des dénominations a pu donner naissance à un ou plusieurs dérivés : *peuplaie* / *peupleraie*, *coudraie* / *coudreraie*, *ormaie* / *ormeraie* / *ormelaie*, *boulaie* / *bouleraie*, *fouaie* / *foutelaie*, *sapaie* / *sapinaie*, *teillaie* / *tillaie* / *tilleraie* / afr. *tillolaie*... Certains ont subsisté après la disparition de la forme qui leur a donné naissance. *Sapaie*, donné comme caduc depuis le moyen français par Diekmann (1975 : 133) mais encore attesté au 20<sup>e</sup> siècle,

- (4) D'autre part, il a eu le mérite d'appeler l'attention sur le sapin de Normandie. Il a bien défini les conditions d'existence des « sapaies » spontanées, et il a préconisé l'emploi du sapin pour l'enrichissement des forêts feuillues. (*Revue forestière française*, 1958, p. 578)

[documents.irevues.inist.fr/bitstream/handle/.../RFF\\_1958\\_8-9\\_578.pdf?...1](http://documents.irevues.inist.fr/bitstream/handle/.../RFF_1958_8-9_578.pdf?...1)

n'a pu être construit que sur l'ancien français *sap*. A partir de *sapin*, il n'y aurait eu aucune raison de tronquer la finale (*sapinaie* est d'ailleurs attesté parallèlement, à côté du plus courant *sapinière*). Toutes les formes, cependant, n'ont pas for-

cément une explication historique : *bouleraie* est plus vraisemblablement de résultat d'une dissimilation et de l'attraction de la sous-série en *-eraie* que le dérivé d'un hypothétique *°boulrier*.

L'essentiel est que dans chaque famille dérivationnelle cohabitent ou ont cohabité plusieurs bases possibles d'un côté, plusieurs dérivés de l'autre, sans que chaque dérivé reste forcément en relation avec sa base formelle.

(5) *orme* : *ormeau, ormier* : *ormaie, ormelaie, ormeraie*

N'importe lequel des trois dérivés désignant un collectif peut être rattaché à n'importe lequel des noms d'arbres. Un dialecte donné peut utiliser aujourd'hui *ormeraie* pour le collectif alors que pour l'arbre il emploiera *orme* (ou *ormeau*) plutôt qu'*ormier*.

Peut-on étendre ce type d'explication à des formations plus récentes, dans la mesure où de nouveaux doublets continuent à se créer parmi les noms d'arbres (*supra*, 2.1.3) ? *Bambousier* est attesté, il pourrait fournir le radical de *bambouseraie* comme *bambou* celui de *bambousaie*. Même si *bambousier* est beaucoup moins diffusé que *bambou*, il serait plausible que la langue sélectionne dans la famille dérivationnelle, parmi les thèmes disponibles, la forme qui assure une meilleure intégration dans la série dérivationnelle du dérivé. Deux types de difficultés s'opposent à la généralisation de cette analyse :

- Certains noms de collectifs en *-eraie* sont sans doute antérieurs au nom d'arbre en *-ier* correspondant, ou étrangers à la communauté où celui-ci est apparu. C'est vraisemblablement le cas de *teckeraie* par rapport à *teckier*, par exemple. Pour construire systématiquement les dérivés en *-eraie* sur le nom d'arbre en *-ier*, il faudrait multiplier les formes virtuelles (*°mélézier*, *°eucalyptier*, *°pruch(i)er*, *°douglas(s)ier*, etc. pour *mélèzeraie*, *eucalypteraie*, *prucheraie*, *douglas(s)eraie...*) au-delà du raisonnable.
- Pour qu'un nom de végétal intègre le paradigme des dérivés en *-ier*, il faut qu'il désigne un arbre prototypique (*palétuvier*), un arbuste donnant une production intéressante (*rosier*) ou au minimum une plante donnant des fruits (*fraisier*). Ce n'est pas le cas pour un certain nombre de bases de collectifs en *-eraie* (*algueraie*, *aspergeraie*, *ignameraie*, *iriserai*, *lavanderaie*, *sisaleraie*, *tuliperaie...*). On ne peut pas avoir *\*irisier* pour la plante qui donne la fleur appelée *iris*.

Les deux séries (ou sous-séries) dérivationnelles, celle des noms d'arbres en *-ier* et celle des noms de plantation en *-eraie*, se développent parallèlement, mais pas à la même vitesse : la première s'étend à la marge, la seconde tend à devenir hégémonique en dehors du discours spécialisé. Elles sont évidemment liées, mais la seconde ne dépend pas de la première. On ne peut donc pas faire l'économie de la variante *-eraie*.

Cas particulier qui fait la transition avec le cas de figure suivant : celui du doublet *olivaie* / *oliveraie*. On sait que le français classique avait gardé quelques traces du système (conservé par l'italien) dans lequel le nom de l'arbre peut être homophone du nom du fruit mais s'en distinguer par le genre : *une olive* 'fruit' / *un olive* 'arbre', *une*

*orange* ‘fruit’ / *un orange* ‘arbre’. *Olivaie* a plus vraisemblablement été construit sur *olive* ‘arbre’ que sur *olive* ‘fruit’ (tandis qu’*oliveraie* est construit sur *olivier*)<sup>9</sup>. *Orangeaie* est plus rare mais également attesté.

### 2.2.3. *Le nom de la plante est construit sur le nom du fruit*

L’arbre et le fruit sont deux entités bien distinctes. Assez proches, cependant, pour que dans certaines langues, ils soient, pour une même espèce, désignés par un même mot (cf. Koch 1999). En français, quand il s’agit d’arbres prototypiques et de fruits au sens courant du terme, peu échappent à la série des dérivés en *-ier*. Sauf exception<sup>10</sup>, le schéma canonique fruit / arbre / plantation s’impose alors comme en (6a).

(6a) *châtaigne* : *châtaignier* : *châtaigneraie*

Quand il s’agit d’arbustes, *a fortiori* de plantes herbacées, les choses sont beaucoup moins nettes. *Planter des cerises* serait incongru mais *planter des fraises* est courant, même si *planter des fraisiers* l’est davantage. *Planter des cassis* est plus fréquent que *planter des cassissiers*. Le paradigme a été étendu à des dérivés dont la base désigne une autre production (*rosier*, *cotonnier*, *caféier* ...), mais à partir d’un nom de fleur, par exemple, le type *rose* / *rosier* est une exception : *camélia*, *azalée*, *rhododendron* désignent aussi bien la plante que la fleur.

Pour la construction du nom de collectif, par conséquent, il n’est pas étonnant que le nom de fruit puisse se substituer au nom de plante quand celui-ci est rare (*kiwitier*) et même quand il est courant (*fraisier*, *cerisier*).

(6b) *kiwi* : (*kiwitier*) : *kiwiraie*  
*fraise* : *fraisier* : (*fraiseie*) / *fraiseraie*  
*cerise* : *cerisier* : *cerisaie* / (*ceriseraie*)

Le locuteur qui forge *kiwiraie*

(7) Voilà la deuxième année que j’achète des pieds de kiwis [...] Je voudrais faire une *kiwiraie* d’une cinquantaine de pieds, 5 rangs palissés de 10 pieds.  
[www.greffer.net/forum/viewtopic.php?t=399](http://www.greffer.net/forum/viewtopic.php?t=399)

a en tête *pieds de kiwis* plutôt que *actinidias* ou *kiwitiers*. Celui qui produit *fraiseie*

<sup>9</sup> Le *FEW* (VII : 349a) considère *olivaie* comme une formation latine, un descendant du pluriel de *olivetum*. Vraisemblable pour l’occitan *oliveta* (ou au masculin pour l’italien *oliveto*), cette paternité l’est moins pour le français *olivaie*, apparu au 17<sup>e</sup> siècle seulement. Nyrop (III : 89) explique de la même façon *prunaie* et *cerisaie* mais nous n’avons pas trouvé trace de °*prunetum* ou de °*cerasetum*.

<sup>10</sup> Officiellement, l’arbre qui donne des kakis est appelé *plaqueminier*, mais *kaki* est beaucoup plus courant. *Kakier*, *kakitier*, *kakissier*, *kakirier* sont attestés mais peu employés.

- (8) Je veux refaire une petite « fraisaie », et j'aimerais trouver une variété bien sucrée et parfumée. Quelqu'un a une suggestion ? Merci ! ...  
[gazette-des-jardins.les-forums.com/.../StolonsDeFraisiers](http://gazette-des-jardins.les-forums.com/.../StolonsDeFraisiers)

ne connaît pas *fraiseraie*, vraisemblablement, puisqu'il éprouve le besoin de forger *fraisaie*, qu'il met entre guillemets. Mais il connaît certainement *fraisier*. Si *fraisaie* lui vient spontanément, plutôt que *fraiseraie*, c'est sans doute que les fraises l'intéressent plus que les fraisiers. Au-delà des considérations pragmatiques, cependant, le court-circuit qui rattache *fraisaie* directement à *fraise* n'empêche pas son sens référentiel de rester 'plantation de fraisiers'. L'essentiel, d'un point de vue constructionnel, est le fait que le dérivé soit mis en relation simultanément avec plusieurs membres de la famille dérivationnelle.

Il n'est pas impossible, pour *fraisaie* comme pour *cerisaie*, *mûraie*, *poiraie*, *prunaie*, *fruitaie*, *mandarinaie*, qu'une contrainte dissimilative intervienne également (*supra* 1.2). Mais elle ne serait pas dans ce cas particulièrement pressante (les deux /r/ de *fraiseraie* sont assez éloignés l'un de l'autre). Et elle n'expliquerait pas *amandaie* (admis par les dictionnaires à côté d'*amanderaie*) ou *arganaie*, *argousaie*, *citronnaie*, *cornouillaie*, *cotonnaie*, *figuaie*, *framboisaie*, *grenadaie*, que nous avons rencontrés également, tous doublets de dérivés en *-eraie*. Le plus vraisemblable, pour *cerisaie* et ses semblables, est que la substitution paradigmatique offre une solution à une difficulté d'ordre phonologique.

Autre difficulté phonologique : les cas où, dans la base, le suffixe *-ier* est précédé d'une voyelle (*noyer*, *caféier*, *cacaoyer*, *théier*). L'allomorphie /je/ ~ /əɾ/ y donne une séquence voyelle + yod + chwa, dans laquelle l'amuïssement du chwa entraîne celui du yod. *Noyer* → *noyeraie* évolue normalement vers *noi(e)raie*, comme *voyer* donne *voirie* et *écurier* *écurie*. Toutes formes où la base immédiate est peu reconnaissable. *Noi(e)raie* renvoie directement à *noix*, comme, d'une autre façon, *noiseraie* qui l'a remplacé. Sur *caféier*, *cacaoyer*, *théier*, aucune forme ne s'est imposée : on trouve *caféieraie*, *cacaoyeraie*, *théieraie*, qui semblent surtout graphiques (comment les prononce-t-on ?), *caféteraie*, *cacaoteraie*, *caféraie*, *cacaoraie*, *théraie*, formés sans doute sur *café*, *cacao* ou *thé*.

Cas limites : celui d'*agrumeraie*, construit sur *agrumes* plutôt que sur le rare *agrumier*, et surtout celui de *glandaie* 'lieu planté de chênes' (*Rob.*). C'est une véritable métonymie qui fait de *glandaie* un synonyme de *chênaie*, puisque dans ce cas le nom du fruit n'appartient pas à la même famille que celui de l'arbre. Mais on voit bien pourquoi elle a été opérée : le 'lieu planté de chênes' sera une *chênaie* si l'on s'intéresse surtout aux arbres eux-mêmes, une *glandaie* si l'on veut y mener des porcs pour qu'ils se repaissent de leurs glands. Indirectement, ce cas éclaire celui de *fraisaie* et des exemples semblables : la substitution du nom du fruit à celui de l'arbre est, pour une part, une forme de métonymie.

#### 2.2.4. Le radical est influencé par un autre membre de la famille dérivationnelle

Il arrive enfin que la formation du radical mette à contribution un élément plus éloigné de la famille dérivationnelle. *Prunelaie* 'verger de pruniers' est facilité par le fait que, dans la famille de *prune* et de *prunier* il y a aussi *prunelle*, même s'il ne s'agit



pas du même fruit et du même arbre. De la même façon, *noiseraie* désigne tantôt un ‘lieu planté de noyers’, tantôt un ‘lieu planté de noisetiers’, par glissement d’un arbre à l’autre dans la même famille (dérivationnelle sinon botanique). *Herbaçaie* ne peut être construit formellement que sur l’adjectif *herbacé*. Pour le *robinier*, on trouve *robinaie* (construit régulièrement sur le nom français) et *robiniae* (sur le latin *robinia pseudoacacia*), mais aussi *robinaie*, dont le radical suppose de remonter à *Robin*, le jardinier qui a donné son nom à l’arbre.

Une place particulière doit être faite aux dérivés en *-ier* (autres que les noms d’arbres ou d’arbustes) appartenant à la même famille, qu’il s’agisse d’adjectifs, de noms de personnes ou autres. L’ancien français *jardinaie* a sans doute emprunté son radical à *jardinier* ; *ronceraie* a été facilité par *roncier* ‘buisson de ronces’ ; *pineraie* par (*pin*) *pinier* ‘pin parasol’ ; *alfateraie*, *endiveraie*, *luzernaie*, *safranaie*, *tomateraie* par *alfatier*, *endivier*, *luzernier*, *safranier*, *tomatier* ‘producteur de N’.

### 3. UNE RÈGLE ? PLUSIEURS RÈGLES ? OU PAS DE RÈGLE ?

Si l’on compare les formations récentes au corpus des dictionnaires, on constate que la dérivation en *-aie* est restée homogène sur le plan sémantique (il s’agit toujours de former des noms de collectifs de végétaux<sup>11</sup>) mais qu’elle s’est diversifiée. Dans la série dérivationnelle, trois ensembles se dessinent assez nettement :

- (1) Une sous-série traditionnelle, qui réunit les dérivés anciens mais continue à s’étendre. La base y est un nom d’arbre ou, s’il s’agit d’un autre végétal, c’est un nom en *-ier* ; le dérivé désigne plus souvent une plantation qu’un milieu naturel (*cannaie*, *jonchaie*, *ronceraie* font exception de ce double point de vue). Le radical est un thème populaire. Le suffixe a le plus souvent la forme standard *-aie* (sauf exception, les finales en *-eraie* résultent de l’allomorphie d’une finale en *-ier* du côté de la base).
- (2) Une sous-série innovante. La base désigne n’importe quel végétal et le dérivé un lieu de culture (*citrouilleraie*, *endiveraie*, *irisaie*, *jasmineraie*, *patateraie*, *pimenteraie*, *pivoine-raie*, *safranaie*, *tomateraie*, *tuliperaie*...) <sup>12</sup>. Ou bien, s’il s’agit d’une végétation spontanée, le dérivé est en général une formation plus ou moins ludique (*chardonneraie*, *coquelicoteraie*, *margueriteraie*, *pissenliteraie*...). Le radical est un thème populaire. Le suffixe a très majoritairement la forme *-eraie*.

<sup>11</sup> Exception (ancienne) dans le corpus provenant des dictionnaires : *tangaie* (ou *tanguaie*) ‘lieu couvert de tangué’ (Rob.). Nous n’avons pas fait de recherches sur la Toile pour savoir s’il y en a d’autres semblables.

<sup>12</sup> Notons en passant, puisque ce n’est pas l’objet de la présente étude, que *-(er)aie* se diffuse ainsi aux dépens de *-ière*. Le double critère que Tomassone et Combettes (1970) avaient observé pour la répartition de *-aie* vs *-ière* (groupement naturel vs artificiel, arbres ou arbustes vs autres plantes) est de moins en moins valable. *Artichauteraie*, *champignonneraie*, *cressonnaie*, *houblonnaie*, *luzernaie*... concurrencent *artichautière*, *champignonnière*, *cressonnière*, *houblonnière*, *luzernière*...

- (3) Une sous-série « savante », forgée par les naturalistes. La base désigne n'importe quel végétal et le dérivé une formation végétale naturelle. Le radical est emprunté au nom latin de la plante. Le suffixe a presque toujours la forme standard *-aie*.

Peut-on ranger toutes ces dérivations sous une même règle ? D'un point de vue référentiel, le résultat de la dérivation reste le même. Mais pas la sémantique de la dérivation elle-même (le nom du fruit peut se substituer à celui de l'arbre) ni, surtout, les données morphophonologiques (sélection du thème, forme du suffixe). La notion de règle impliquant celle d'exception, il faudrait multiplier les exceptions au-delà du raisonnable.

Formuler une règle pour chacune des trois sous-séries ? La troisième pourrait être mise facilement à part, si l'on décide qu'elle appartient à une langue de spécialité. Ce qui est net pour les discours où l'on trouve les formations concernées. Mais elles y côtoient les formations populaires. On trouve dans les mêmes contextes tantôt *érbiaie*, tantôt *acéraie*, tantôt *populaie* associé à *saulaie*, tantôt *peupleraie* à côté de *salichaie*.

- (10a) En s'éloignant du cours d'eau, on passe de la populaie-saulaie à peuplier noir, saule pourpre [...] à la populaie-saulaie à peuplier noir, saule blanc [...]  
[basecommunale.paca.ecologie.gouv.fr/pdf/fiches/znieff2.../84128100.pdf](http://basecommunale.paca.ecologie.gouv.fr/pdf/fiches/znieff2.../84128100.pdf)
- (10b) [...] peupleraies blanches et autres forêts dites à « bois dur » des régions nord-méditerranéennes [...] quelques relevés de Salichaies (*Salix pedicellata*) [...]  
[ecologia-mediterranea.univ-avignon.fr/.../Ecologia\\_mediterranea\\_1999-25\\_1\\_01.pdf](http://ecologia-mediterranea.univ-avignon.fr/.../Ecologia_mediterranea_1999-25_1_01.pdf)

Les interférences sont nombreuses, dans un sens comme dans l'autre. *Tauzaie* est un compromis entre ce que donneraient une dérivation populaire (*°tauzinaie* ← (*chêne*) *tauzin*) et une dérivation savante (*tozzaie* ← (*quercus*) *tozza*). *Querceraie* combine une base latine (comme les formations spécialisées) et la variante *-eraie* (comme celles du lexique général) pour désigner non pas un biotope mais une collection de chênes.

- (11) Visite de la querceraie (170 variétés de chênes), du potager jaune en fleur, des sculptures réalisées après la tempête de 1999 par les artistes [...]  
[www.poitou-charentes.culture.gouv.fr/v4.htm/.../jep09\\_charente.pdf](http://www.poitou-charentes.culture.gouv.fr/v4.htm/.../jep09_charente.pdf)

Certains doublets correspondent à une répartition des rôles (*pistaciaie* désigne un biotope naturel, *pistacheraie* une plantation d'arbres fruitiers), mais pas tous : *cistaie* et *cisteraie*, *eucalyptaie* et *eucalypteraie*, apparaissent dans des contextes semblables. Les trois sous-séries interfèrent largement, formuler trois règles ne diminuerait pas le nombre des exceptions.

Mais faut-il absolument des règles ? On doit rendre compte du fait que la synchronie actuelle puisse produire à la fois *fraisaiie* (alors qu'il s'agit de fraisiers et pas de fraises) et *endiveraie* (alors que l'endivier n'est pas la plante qui donne les endives mais celui qui les cultive) à côté ou à la place de *fraiseraie* et de *°endivaie*. Difficile pour une règle de prévoir l'un et l'autre. On peut, en revanche, concevoir la formation

des mots comme une dynamique proprement lexicale où chaque dérivation s'opère au croisement de deux paradigmes, une série dérivationnelle et une famille dérivationnelle, et sous l'influence de contraintes plus ou moins universelles de « bonne formation ». Dans ce cadre, rien ne s'oppose à ce qu'un même modèle dérivationnel donne des formes différentes à partir d'une même base, si toutes ces contraintes (lexicales ou phonologiques) sont transgressables et si leur hiérarchisation est variable (en fonction de l'époque, du type de discours, des données pragmatiques, etc.). Or c'est précisément ce que montre le corpus, où les variations sont particulièrement nombreuses (pour plus d'un tiers des bases, on trouve plusieurs formes du côté du dérivé). La pression de la série dérivationnelle, la tendance à généraliser les finales en *-eraie*, la présence d'*endivier* dans le même type de contexte (ce sont les endiviers qui gèrent les endiveraies) font préférer *endiveraie* à *°endivaie*. La possibilité de substituer un radical à un autre, à l'intérieur de la famille dérivationnelle, permet *fraisais* malgré la contrainte de fidélité, la contrainte prosodique et la pression de la série dérivationnelle qui ont généralisé *fraiseraie* sans empêcher totalement une forme concurrente.

## CONCLUSION

Les enseignements qu'on peut tirer de cette étude sont de deux ordres. Sur les faits linguistiques eux-mêmes, elle révèle une créativité lexicale, un foisonnement de formes insoupçonnables à partir des intuitions du linguiste, même complétées par les ressources des dictionnaires. La dérivation en *-aie*, tout en restant circonscrite à un domaine bien particulier, est beaucoup plus productive et diversifiée qu'on aurait pu le croire, et les aléas morphophonologiques ne s'y limitent pas à l'application d'une règle d'allomorphie contrariée par quelques bizarreries ponctuelles. Sur la méthode, cette recherche confirme non seulement l'intérêt mais la nécessité de rassembler des bases de données extensives, qu'il s'agisse de l'héritage historique toujours vivant ou des productions contemporaines auxquelles la Toile permet d'avoir accès. Elle confirme également la nécessité de croiser les approches. Les paramètres proprement morphologiques sont inséparables de la dimension lexicale (dynamique de la série dérivationnelle, ressources offertes par la famille dérivationnelle) et des contraintes phonologiques, sans oublier l'éclairage historique. Les différentes explications sont souvent redondantes, et dans ce cas il est difficile de choisir (à supposer qu'il faille choisir). Elles sont en fait complémentaires et l'on a besoin de toutes. *Cerisais* peut se lire comme une contraction de *ceriseraie*, pour obéir à la contrainte dissimilative, ou comme le résultat d'une substitution de *cerise* à *cerisier*. On peut voir dans *joncherais* l'interfixe *-er-* ou la variante *-eraie*. *Boulaie* a sans doute été construit sur l'afr. *boul* et *noiseraie* sur *noisier* (doublet ancien de *noyer*), mais ils sont aujourd'hui rattachés à *bouleau* et à *noix*, et il faut rendre compte autrement des écarts formel ou sémantique que cela crée. Et même si *boulaie* est bien installé, l'apparition de *bouleaurais* révèle que la contrainte de fidélité et l'attraction de la série des dérivés en *-(e)raie* peuvent toujours susciter une forme nouvelle.

## Références

- ALEXANDER, Luther Herbert (1912) *Participial Substantives of the -ata Type in the Romance Languages*. New-York : The Columbia University Press.
- ARONOFF, Mark (1994) *Morphology by Itself. Stems and Inflectional Classes*. Cambridge (Mass.) : The MIT Press.
- BONAMI, Olivier/Gilles BOYÉ (2005) « Construire le paradigme d'un adjectif ». *Recherches linguistiques de Vincennes* 34, 77-98.
- CORBIN, Danielle (1987) *Morphologie dérivationnelle et structuration du lexique*. Tübingen : Niemeyer.
- DIEKMANN, Erwin (1975) « Wortbildungsstruktur und Mengenlehre. II. Bezeichnungen für den Ort, wo Pflanzen wachsen, auf *-aie* bzw. *-ière* ». *Zeitschrift für romanische Philologie* 91, 129-142.
- FEW = *Französisches etymologisches Wörterbuch*, de W. von Wartburg et continuateurs. Bonn, puis Bâle : Helbing und Lichtenbahn, 1928-.
- KOCH, Peter (1999) « TREE and FRUIT: a cognitive-onomasiological approach ». *Studi Italiani di Linguistica Teorica ed Applicata* 28, 331-347.
- PLÉNAT, Marc (2009) « Les contraintes de taille ». In : B. Fradin/F. Kerleroux/M. Plénat (dir.), *Aperçus de morphologie du français*. Saint-Denis : Presses Universitaires de Vincennes, 47-63.
- PLÉNAT, Marc (2011) « Enquête sur divers effets des contraintes dissimilatives en français ». In : M. Roché/G. Boyé/N. Hathout/S. Lignon/M. Plénat, *Des unités morphologiques au lexique*. Paris/Londres : Hermès-Lavoisier, 145-190.
- PLÉNAT, Marc/Michel ROCHÉ (2004) « Entre morphologie et phonologie : la suffixation décalée ». *Lexique* 16, 159-198.
- Rob.* = *Le Grand Robert de la Langue Française, Dictionnaire alphabétique et analogique de la langue française* de Paul Robert, 2<sup>e</sup> édition revue par Alain Rey. Paris : Le Robert, 1987.
- ROCHÉ, Michel (2003) « L'interfixe est-il une unité morphologique ? ». In : B. Fradin et al. (éds) *Les unités morphologiques*, Actes du 3<sup>e</sup> Forum International de Morphologie (Villeneuve d'Ascq, 19-21 septembre 2002), *Silicales* 3, 169-178.
- ROCHÉ, Michel (2008) « Structuration du lexique et principe d'économie : le cas des ethniques ». In : J. Durand/B. Habert/B. Laks (éds), *Congrès Mondial de Linguistique Française - CMLF'08*. Paris : ILF, 1571-1585.
- ROCHÉ, Michel (2010) « Base, thème, radical ». *Recherches Linguistiques de Vincennes* 39, 95-133.
- ROCHÉ, Michel (2011) « Quel traitement unifié pour les dérivations en *-isme* et en *-iste* ? ». In : M. Roché/G. Boyé/N. Hathout/S. Lignon/M. Plénat, *Des unités morphologiques au lexique*. Paris/Londres : Hermès-Lavoisier, 69-143.
- TOMASSONE, Roberte/Bernard COMBETTES (1970) « Remarques sur la formation des collectifs de noms d'arbres et de plantes ». *Revue de Linguistique Romane* 34, 224-233.

TLF = *Trésor de la langue française, Dictionnaire de la langue du XIX<sup>e</sup> et du XX<sup>e</sup> siècles*, sous la direction de P. Imbs et de B. Quemada. Paris : Editions du CNRS (t. 1-10) et Gallimard (t. 11-16), 1971–1994.

### Résumé

#### PRESSION LEXICALE ET CONTRAINTES PHONOLOGIQUES DANS LA DÉRIVATION EN *-AIE* DU FRANÇAIS

La dérivation en *-aie* (*chênaie*, *roseraie*, *châtaigneraie*...) suit normalement le schéma : N de plante + *-aie* → N de collectif. Quelques dérivés présents dans les dictionnaires, cependant, s'écartent de cette configuration : *cerisaie* au lieu de *ceriseraie*, *joncheraie* à côté de *jonchaie*, etc. Ils sont beaucoup plus nombreux, proportionnellement, parmi ceux qu'on peut recueillir sur la Toile. On a mis en évidence deux séries de facteurs, tantôt convergents tantôt antagonistes, qui expliquent ces écarts : (i) l'influence de la série dérivationnelle dans laquelle le dérivé est destiné à s'intégrer (un « effet de rime » tend à substituer à la finale phonologiquement « pauvre » *-aie* la finale plus « riche » *-eraie*, déjà surreprésentée à cause du grand nombre de noms d'arbres en *-ier* parmi les bases des dérivations antérieures) ; (ii) des contraintes de « bonne formation » – contraintes de taille, contraintes dissimilatives – qui conduisent à modifier le thème fourni par la base afin qu'il constitue un radical phonologiquement plus satisfaisant. Le corpus révèle également un grand nombre de dérivés construits sur des thèmes « savants » (*bétulaie*, *lariçaie*, *suberaie*...), bien que le suffixe appartienne lui-même au versant « populaire » de la dérivation ; ce qui conduit à remettre en cause le principe de correspondance univoque entre les cases de l'espace thématique d'un lexème et celles de son paradigme dérivationnel. Ces observations relativisent fortement la notion de « Règle de Construction des Lexèmes » au profit d'un système de contraintes d'ordre lexical ou morphophonologique.

### Povzetek

#### LEKSIKALNI VPLIV IN FONOLOŠKE OMEJITVE V IZPELJAVI S PRIPONO *-AIE* V FRANCOŠČINI

Izpeljava s pripono *-aie* (*chênaie*, *roseraie*, *châtaigneraie*...) ustreza običajno naslednjemu besedotvornemu vzorcu: samostalnik, ki označuje rastlino + *-aie* → kolektivni samostalnik. Nekaj izpeljank, ki jih najdemo v slovarjih, odstopa od te sheme : *cerisaie* namesto *ceriseraie*, *joncheraie* poleg *jonchaie* ipd. Sorazmerno številnejši so primeri, ki jih najdemo na medmrežju. Izpostavili smo dve vrsti dejavnikov, ki delujejo včasih v isti smeri, včasih pa v nasprotnih smereh, in s pomočjo katerih lahko te razlike razložimo: (i) vpliv besedotvorne paradigme, v katero bo vključena izpeljanka (zaradi »učinka rime« bo fonološko »šibkejše« obrazilo *-aie* nadomeščeno z »močnejšim« *-eraie*, ki zaradi velikega števila imen dreves s pripono *-ier*, ki služijo kot podstava te vrste izpeljank, že sedaj prevladuje; (ii) omejitve, ki jih nalaga »pravilno izpeljevanje« (omejitve glede dolžine, razlikovalnosti) in ki privedejo do

sprememb osnove, kakršno ponuja podstava, da bi tako prišli do fonološko ustrežnejše osnove. Korpus prinaša tudi veliko število izpeljank iz »učenih«, latinsko-grških podstav (*bétulaie, lariçaie, suberaie...*), čeprav sodi sama pripona bolj v »ljudski« del izpeljevanja. To postavi pod vprašaj načelo enoznačne ustreznosti med tipom paradigme osnove in tipom besedotvorne paradigme. Vsa omenjena opažanja omajajo splošno veljavnost samega pojma »pravila tvorjenja leksemov«, ki ga kaže dopolniti oziroma nadomestiti s sistemom leksikalnih in morfofonoloških omejitev.

## WORD-FORMATIONAL PRODUCTIVITY OF THE SLOVENE LANGUAGE IN THE CASE OF SPORTS NEOLOGISMS

### 1. INTRODUCTION

1.0 At the turn of the 21<sup>st</sup> century, the world has been undergoing enormous changes reflecting the current global trends towards an amalgamation of local economic, cultural and social structures into broader multinational conglomerates. The main goal of these conglomerates is to enable the flow of capital, ideas and products finding their way into our everyday life (Vuković 2011: 19). Heavily promoted by advertising, these new products produce new lifestyle trends, such as having a successful career, practicing functional use of modern technology and using safe and ecological means of transportation, but above all, modern society values an attractive physical appearance based on the achievement of good physical form.

1.1 These lifestyle trends reflect those areas which have been developing very rapidly, also linguistically, because they have some kind of impact on the general population. These areas include advertising, mobile phone communication, electronic communication, fashion, cosmetics, the car industry, cuisine and sport. In the case of the Slovene language, studies have already been conducted on the use of complex words in the car industry (Bokal 1998), mobile phone communication (Stramljič Breznik 2010) and advertising (Stramljič Breznik, Voršič 2009), so the present study focuses on the field of sports.

### 2. METHODOLOGY

2.0 The article deals with the word-formational characteristics of lexical items that belong to the most general semantic field of sports, further specified by the hypernym *the type of sport* (based on the number of participants). This yields a sample of controllable size containing lexemes with comparable semes which already belong to the general body of Slovene lexis (e.g. according to the *Slovar slovenskega knjižnega jezika*/Dictionary of the Slovene standard language (hereafter - SSKJ): football is 'a sports game in which players try to kick a ball into a goal') as well as the terminolog-

---

\* *Author's address:* Filozofska fakulteta v Mariboru, Oddelek za slovanske jezike in književnosti, Koroška cesta 140, 2000 Maribor, Slovenia. Email: irena.stramljic@uni-mb.si

\*\* *Author's address:* Filozofska fakulteta v Mariboru, Oddelek za slovanske jezike in književnosti, Koroška cesta 140, 2000 Maribor, Slovenia. Email: ines.vorsic@uni-mb.si

ical part of Slovene lexis (futsal ‘a variant of football played by only seven players on a relatively smaller pitch’). The selected lexical items are then analyzed from the perspective of word-formation. Special attention is given to those complex words from the selected word families that represent more recent and word-formationally less established lexis, i.e. words with the status of neologisms and potential complex words as well as nonce words.

**2.1.** The data on the selected lexis was collected from the Slovene reference corpus of texts FidaPLUS with the help of the multifunctional tool Sketch Engine (abbr. SkE). This tool enables classic functions for concordances (search for lemmas, word forms, phrases, search with the CQL language) and standard forms of presenting, classifying, filtering and saving concordances. Additionally, it enables other types of analysis, such as the creation of the Word List, the use of Thesaurus and Word Sketch, as well as Sketch-Diff. In view of these possibilities, the corpus approach is suitable for the analysis of established as well as less established or not yet established lexical characteristics of language, which in our particular case refers to the field of sports. Our research relies heavily on the corpus function of the world list, which enables the extraction of all the word forms or lemmas in the corpus and ranks them by frequency, while also enabling the extraction of groups of words which stem from the same root. Our study is based on the words which were extracted by searching for the root of the word (e.g. \*ples.\*), yielding the word family of the lexeme *ples*. The analysis only included those words that occurred in ten tokens at the most (which was only the case in two examples, the majority had a much lower count), because our focus was directed at the analysis of recent lexis with low token counts. In lexicography, what counts as relevant data is the occurrence of a lexeme in three different independent sources (Stramljič Breznik 2004). In the study presented here we were interested in words which have not yet established themselves and which are still making their way into the lexical inventory of the Slovene language. For this reason we followed the criterion according to which the token does not appear in more than in three different sources. Lexical units that were the result of typing mistakes were also excluded.

### **3 DEFINITION OF BASIC LEXICOLOGICAL CONCEPTS**

**3.0** Methodological framework contains several lexicological concepts, such as the semantic field, word family, complex word, neologism, potential complex word and nonce word. For the purpose of clarity these concepts have to be given a general definition and illustrated with concrete examples.

**3.1.** In order to achieve an objective selection of lexemes from the field of sports, we have to keep in mind that the lexical structure of language is systemic. According to Jost Trier (1931), who is considered the founder of the theory of systemic lexical structure, the whole lexicon of a language works as a structure of interconnected fields which are organized in a hierarchical and paradigmatic way, and these fields themselves have a structure consisting of conceptually related words.



In Slovene linguistics, the overview of the concepts and terminology used for naming semantically organized groups of words by Stramljič Breznik (2007: 97 – 109) has attracted a high level of interest from numerous authors, inciting a discussion involving different positions. In Slovene linguistics, just as in other countries (Nikolić-Hoyt 2004: 109), the concept of the lexical organization of language originated primarily from practical considerations of lexicography and not from theoretical considerations of semantics. In the year 2000 a number of conceptualizations of possible terminological groups emerged, which resulted in several versions of the terminology. In *Slovenska slovnica/Slovene Grammar* (Toporišič 2000: 118, 1996: 89) the term word field (*besedno polje*) is used, defined as a group of closely connected words that name extra-linguistic reality by sharing the same spatial, temporal or causal frameworks. In their textbook, Križaj Ortar, Bešter Turk et al. (2000: 13 -18) use the term thematic field, defining it as a group of words whose use in a text can be predicted solely on the basis of the topic or the title of the text. The third term, which was introduced by Vidovič Muha (2000: 59), is semantic field, defined as a conceptual field structured by semantic constituents, i.e. cognitive reality formed by the classifying semantic constituent.

In order to find an appropriate selection of lexemes for the purposes of this research we had to take into account the concept of semantic field, which was at first specified as the most general semantic field of sports. Within the general field, three semantically more specific semantic fields were determined: first, team sports with a selection of basic lexemes: football/footy ; handball, volleyball, basketball, ice hockey, baseball, rugby, etc; second, individual sports: athletics, gymnastics, cycling, motor sports, tennis, wrestling, boxing, karate, fencing, swimming, sailing, skiing, sledding, ice skating, surfing, golf, rollerblading, body building, etc. Potentially, there is a third semantic field, which can be defined by the hypernym pair sports, such as dancing, rowing (e.g. coxless pair). Here it has to be noted that the structuring of the semantic fields solely on the basis of the number of participants, while appropriate for the purpose of this research, does not exclude other various possible classifications of varieties of sports on the basis of other criteria, such as the space in which it is carried out (e.g. indoor v. outdoor sports), the object used in the sport (ball, puck, target, bicycle, etc.), skills involved (fighting v. non-fighting sports), the way it unfolds (water sports, air sports), the season in which it is carried out (winter v. summer sports), etc.

The selection of lexemes from individual semantic fields was implemented on the basis of two criteria. The first criterion was the visibility and popularity of the sport in Slovenia, e.g. semantic fields such as those of baseball and rugby were excluded. The second criterion included the number of motivated words. In others words, the selection included those lexemes from which a greater number of complex words were formed, and which thus presented the basis of a larger word family.

**3. 2.** A complex word is a motivated word which is semantically and formally related to its motivating word at the synchronic level. It is a bipartite word, consisting of a base and an affix or affixes. A group of complex words sharing the same root is

called a word family (Toporišič 1992: 7; 1996: 89; 2000: 108, 116); Vidovič Muha 2000: 107 – 108); Stramljič Breznik 2004).

In our particular case, the inclusion of all the members of the word family of a selected lexeme from a particular semantic field would miss the purpose of the study, because our interest was in more recent, not yet established expressions, which are characterized by a low count. The complex words selected thus had the status of neologisms, more specifically the status of a potential word or a nonce word.

3.3. Neologism is a term which is mostly used to refer to new words. Such use of the term, however, is only valid in the case of new expressions created for a new concept. For this reason, the most appropriate definition of neologisms is the one which is based on the place of new lexemes in the existing lexical structure. Neologisms can thus be defined as those words which have the potential to become a unit of language, i.e. a potential to enter the lexical system of the language (Martincová 1972). On the basis of necessity, neologisms are divided into necessary and stylistic neologisms. The former are new expressions created for naming new objects or phenomena, i.e. filling a lexical gap, while the latter are ad-hoc words created for a specific occasion (Muhvič-Dimanovski 2005: 6 – 7). Stylistic neologisms are formed in parallel to already existing expressions and are mostly context-dependent and therefore expressive.

In Slovene linguistic literature, a neologism is defined as a new word (*Slovenski pravopis/Slovene Orthography*), and as a not yet generally established word or phrase (SSKJ, 1998). Toporišič (1992: 2001) defines a neologism as a newly formed word or phrase or an old word or phrase used in a new way, explaining it as a recent formation, a temporally marked word which can become a part of neutral vocabulary through frequent use, otherwise remaining an ad hoc formation (Toporišič 2000: 130). These definitions imply a distinction between a new word (a neologism) and an ad hoc word (a nonce word). Nonce words can be defined as words created for a particular occasion, as words which “the speaker forms according to regular word-formational patterns and are understood by the hearer or reader. But they do not become part of general usage” (Toporišič 2000: 130 – 131), and are not institutionalized because they are limited to particular speech or textual situation (Toporišič 1992: 222).

By taking into account the corpus data about the occurrence and the systematic nature of their formation, the study included those neologisms that fall into the following three categories. The first category comprises complex words which are formed according to the existing word-formational patterns and have the potential to enter the lexical system, while the second category comprises complex words which are formed according to the regular word-formational patterns, but have a marked combination of bases and/or affixes. The third category consists of those complex words which are very individual and are not created according to regular patterns or do not have a predictable morphemic combination. According to this classification, only the complex words from the first category have the status of neologisms as potential lexical units, while the second and the third categories deviate from that by entering the sphere of individual and ad-hoc usage, showing creativity in language use and intentional expressiveness.

## 4 SELECTION OF MATERIAL

4.0 On the basis of the above-mentioned methodological framework and according to the theoretical definition of the key terms specified for the purpose of this research, a selection of the following base words (i.e. motivating words) yielding low frequency complex words which have not yet established themselves as part of vocabulary have been selected.

### 4.1 Sports

The word family with the base word *šport* (sport) includes the following lexemes:

*aerošport* (1), *antišportnica* (1), *antišportnik* (1), *bikešport* (2), *kvazišportnica* (1), *kvazišportnik* (1), *nešport* (1), *neošportnik* (1), *športoid* (1), *športoljubec* (1), *športolog* (4), *športoslovec* (3), *športologija* (5), *športokracija* (2), *športovščina* (7), *športmen* (1), *športnež* (3), *športščina* (3), *ultrašport* (2), *ultrašportnik* (1), *vsešportnik* (1), *žogošport* (2); *antišporten* (1), *jugošporten* (2), *kvazišporten* (1), *medšporten* (1), *multišporten* (1), *parašporten* (1), *pošporten* (1), *polšporten* (1), *športnoaktiven* (3), *športnojahalen* (1), *športnokomentatorski* (1), *športoholičen* (1), *športološki* (1), *športosloven* (3), *ultrašporten* (1), *velešporten* (1), *všešporten* (1), *zabavnošporten* (1), *zašportan* (1), *žogošporten* (2); *športajoč* (1).

From the semantic field *team sports* the following lexemes yielding members of word families have been selected:

“Nogomet”/“fuzbal”/“nogobrc” (standard and substandard terms for football):

*antinogomet* (3), *antinogometaš* (1), *evronogomet* (1), *jugonogometaš* (1), *nadnogometaš* (1), *najnogometaš* (2), *nogomanija* (1), *nogometafizika* (2), *nogometarjenje* (1), *nogometaška* (1), *nogometnik* (1), *nogometoljubec* (1), *nogometolog* (1), *nogometomanija* (1), *nogometoslovec* (1), *nogometščina* (6), *pranogomet* (1), *slučinogometaš* (1), *supernogometaš* (1), *velenogometaš* (1), *velikonogometaš* (3); *antinogometen* (1), *jugonogometen* (4), *mednogometen* (1), *nadnogometen* (1), *nogometalen* (1), *obnogometen* (7), *protinogometen* (3), *sociološko-nogometološki* (1), *velenogometen* (2), *velikonogometen* (7); *ponogometiti* (1); *fuzbalček* (1), *fuzbalerščina* (1), *fuzbalfen* (2), *fuzbalogija* (1), *fuzbalolog* (1), *fuzbalomanija* (5), *fuzbalščina* (9), *jugofuzbal* (1); *antifuzbalski* (1); *nogobrc* (3), *nogobrcar* (1), *nogogazem* (1), *nogohitrec* (2), *nogoljublje* (1), *nogovanje* (1).

“Košarka”/“basket” (standard and substandard term for basketball):

*evrobasket* (2), *košarkašenje* (3), *košarkiranje* (2), *jugokošarkar* (3), *minikošarka* (3), *najkošarkar* (2), *nekošarkar* (1), *nekošarkaš* (1); *jugokošarkarski* (81), *nadkošarkarski* (1), *obkošarkarski* (1), *vodnokošarkarski* (2); *basketaš* (3), *basketiranje* (2), *minibasket* (4); *basketiški* (1).

“Rokomet” (handball):

*minirokomet* (2), *rokometec* (2); *minirokometen* (1), *obrokometen* (2), *predrokometen* (1), *rokometnonogometnekošarkarskokolesarski* (1).

“Hokeyj” (hockey):

*hokejka* (3), *hokejščina* (4), *nehokejist* (1), *prahokej* (1), *rolerhokej* (1); *hokejaški* (6), *nehokejski* (4), *obhokejski* (1).

“Odbojka” (volleyball):

*miniodbojka* (3), *nogoodbojka* (1); *kvaziodbojkarski* (1); *priodbojkirati* (1).

**4.3** From the semantic field of individual sports the following lexemes yielding suitable family members have been selected:

“Smučanje/skijanje” (skiing):

*e-smučarščina* (3), *parasmučar* (1), *prasmučar* (2), *smučanorija* (2), *smučarajanje* (1), *smučarščina* (2), *smučerator* (1), *smučkice* (1), *supersmučar* (2), *supersmučarka* (2); *alpinističnosmučarski* (2), *predsmučarski* (2), *smučarskorekreacijski* (1), *smučarskotekaški* (10), *smučiščarski* (2), *smučljiv* (6), *velesmučarski* (1); *posmučati* (1); *skijanje* (2),<sup>1</sup> *skijač* (1), *skijaš* (1); *skijaški* (2).

“Atletika” (athletics):

*atletizem* (1), *atletičnost* (1), *atletizacija* (1), *atletizem* (5), *atletskost* (1), *atletstvo* (1), *nadatlet* (1), *najatletinja* (1), *psevdoatletika* (2), *superatletinja* (1), *turboatlet* (1); *atletičen* (2), *superatletski* (1); *atletizirati se* (1).

“Golf” (golf):

*golfač* (6), *golfhec* (1), *golfiat* (1), *golfiga* (2), *golfin* (1), *golfivec* (1), *golfizem* (1), *golfliga* (2), *golfsman* (1), *golfsmanija* (1), *golfsafari* (3), *golfsčina* (1); *antigolfovski* (1), *obgolfski* (2), *pogolfski* (1), *velegolfski* (1); *golffissimo* (1).

“Kolesarstvo” (cycling):

*kolesarijada* (3), *kolesarnina* (1), *nekolesar* (1), *prekolesarjenje* (1), *sokolesar* (2), *ultrakolesar* (5); *postkolesarski* (1), *protikolesarski* (1), *razkolesarjen* (1), *ultrakolesarski* (1), *velekolesarski* (2); *pokolesariti* (5), *obkolesariti* (8).

“Fitness” (fitness):

*akvafitness* (2), *fitnesar* (2), *fitneser* (6), *fitneserka* (1), *fitnesindustrija* (1), *fitnesiranje* (2), *fitnesjoga* (1), *fitnesmojster* (1), *fitnesmuzika* (1), *fitnesrevolucija* (1), *fitnesrutina* (1); *fitneserski* (1), *sfitnesiran* (2); *fitnesirati* (1).

“Tenis” (tennis):

*minitenis* (7), *tenisačenje* (1), *tenisačina* (2), *tenisarka* (10), *tenisiranje* (2), *tenisišče* (1), *teniščina* (3); *tenisaški* (3), *tenisšpilen* (1); *tenisirati* (4).

---

<sup>1</sup> In the corpus this word has 7 occurrences, but in five of these cases it has been used in a text in Croatian.

“Sankanje” (sledding):

*sankalec* (2), *sankarija* (6), *sankarijada* (1), *sankaštvo* (1), *sankijada* (1); *sankarski* (1); *pre-sankati* (1), *zasankati se* (1).

“Kegljanje”/”bovling”/”bowling” (bowling):

*kegljada* (1), *nekegljač* (1); *bovlišče* (4); *bovlinški* (1); *bowlingaš* (1); *bowlati* (1), *pribowlati* (1).

“Boks” (boxing):

*boksaerobika* (1), *boksališče* (1), *boksarščina* (2), *brcoboks* (9), *brcoboksar* (1); *boksarskonazorski* (1).

“Karate” (karate):

*karatejec* (2), *karatejka* (1), *karateseks* (1), *kvazikarate* (2).

**4.4** From the semantic field of *pair sports* only one lexeme with the related word family members has been selected:

“Ples” (dancing):

*antiples* (2), *neples* (2), *plesač* (8), *plesarija* (2), *plesoholičarka* (1), *plesoljubec* (2), *plesovodja* (3), *postples* (2), *soloplesaec* (1); *ognjenoplesen* (1), *plesajoč* (1), *plesaželjen* (1), *plesnobaleten* (1), *plesnodramaturški* (1), *plesnoizrizen* (1), *plesnokoreografski* (1), *plesnomoden* (1), *plesno-ritmičen* (2), *plesolomen* (1), *psevdoplesen* (1).

## 5 WORD-FORMATIONAL ANALYSIS

**5.0** The word-formational analysis is based on the structural-semantic principle. This means that complex words have been analyzed in terms of their structure – in terms of their formation, word category and affixes. The semantic analysis has focused on their word-formational and lexical meaning. On the basis of predictable or unpredictable patterns and morphemic combinations used in their formation, the studied complex words have been classified into the following three categories.

### 5.1 Potential complex words

**5.1.0** These include complex words formed according to predictable word-formational patterns, which means that they are less creative.

#### 5.1.1. Derivatives with suffixation

**5.1.1.1** Derivatives with suffixation include a considerable number of nominal complex words of feminine gender. These make use of the established inventory of Slovene suffixes. Semantically, they form a large group of feminine versions of masculine forms with the following suffixes: *-ica* (*antišportnica*, *kvazišportnica*), *-ka* (*nogomataška*, *supersmučarka*, *fitneserka*, *tenisarka*, *-inja* (*najatletinja*, *superatletinja*)). There is also an interesting use of the generalized suffix *-ščina*, formed on the

analogy with the regular formation using the suffix *-ina* (*slovenščina* < *slovenski jezik*), preserving the palatalized and assimilated part of the base which occurs in such cases. Complex words formed with *-ščina* have a different lexical meaning from those with *-ina*. Whereas words with *-ina* have the meaning of 'language or payment' (e.g. *kolesarnina*/cycling fee), those with *-ščina* form a group of new words used in computer games simulating a variety of sport (*športovščina*, *nogometščina*, *hokejščina*, *šmučarščina*, *golfsčina*, *boksarščina*, *teniščina*; npr. *Za začetek je premikanje igralcev v tej teniščini tako počasno, nadzor pa tako nenatančen, da pri igranju bodisi zehaš, bodisi mečeš n-gage ob tla*). In this semantic group there is one word which differs from the above words in terms of its base, namely *tenisaščina*, which is derived from *tenisač* (compare *teniščina* < *tenis*), and one which is different in terms of its suffix, namely *karatejka*.

A considerable group is comprised of derivatives with the suffix *-ijada* (*kolesarijada*, *sankijada*, *kegljijada*), which remain within a productive and semantically predictable group with the meaning of 'sports event'. The only exception is *sankarijada*, which denotes an activity and is used as a synonym for the dictionary word *sankanje*. Similarly there are derivatives with *-izem*, which retain the meaning of property or state (*atletizem* 'athletic appearance'), but interestingly also as acts (*atletizem* 'athletic training'). The derivative with *-arija* can express a certain degree of positive connotation (*sankarija*), but at the same time this suffix can also have a negative connotation (*plesarija*). New complex words with the suffix *-izacija* also follow a predictable transformational pattern (*atletizacija* < *atletizirati*), using systemic deverbal derivation as in the example *slovenizirati* > *slovenizacija*. Complex words of neuter gender in this category mainly belong to the semantic group denoting the venue of the sports event and are formed with the predictable suffix *-(l)išče* (*tenisišče*, *bovlišče*, *boksališče*).

5.1.1.2 Adjectival derivatives with suffixation also make use of predictable suffixes, e.g. *-ski*, which is added to nominal bases, but in this case the bases are stylistically marked (*smučiščarski* < *smučiščar*, *fitneserski* < *fitneser*, *sankarski* < *sankar*), *-ki* (*športološki* < *športolog*, *hokejaški* < \**hokejač*). The same goes for deverbal derivatives in *-n* (*zašportan* < *zašportati se*, *sfitnesiran* < \**sfitnesirati se*), *-oč* (*športajoč* < *športati*, *plesajoč* instead of *plešoč*). Good use is also made of *-ljiv* with the meaning of inclination to act in the new word *smučljiv*. There is a considerable group of adjectival derivatives which are formed from a prepositional nominal phrase according to the predictable systematic pattern *ta/tak, ki je Prep + N* (he/she who is Prep + N): *tak, ki je ob nogometu* > *obnogometen*. Four prepositions are used in such cases: *ob-* (*obgolfski*, *obhokejski*, *obkošarkarski*, *obnogometen*, *obrokometen*), *po-* (*pogolfski*, *pošporten*), *med-* (*mednogometen*, *medšporten*), and *pred-* (*predrokometen*, *predsmučarski*). The last three of the above-mentioned prepositions can also be interpreted as prefixes transformed from temporal adverbs *potem/kasneje*, *vmes* and *poprej*, in which case they could also be classified as adjectival derivatives with prefixation.

5.1.1.3 Verbal derivatives also follow the predictable denominal pattern with the suffix *-irati* (*fitnesirati*, *tenisirati*). The verbal derivation *bowlati* is particularly interesting as it consists of a foreign base and the native verbal morpheme *-ati*, which makes it a hybrid derivative.

## 5.1.2. Compounds

5.1.2.0 The corpus data confirms the well-known observation that nominal and adjectival compounds are predominant in the Slovene language, whereas verbal compounds are less frequent. According to the typology of lexemes in this category, the group of neo-classical compounds with combining forms (the so-called prefixoid and suffixoid formations) has to be mentioned first. The morphemic structure of these compounds includes truncated forms of generally used international root morphemes. The following examples from the corpus are typical nominal and adjectival prefixoid formations, with nominal and adjectival examples separated by a semi-colon:

*akva-* (*akvafitnes*), *evro-* (*evrobasket*, *evronogomet*), *fitnes-* (*fitnesidustrija*, *fitnesjoga*, *fitnesmojster*, *fitnesmuzika*, *fitnesrevolucija*, *fitnesrutina*), *jugo-* (*jugokošarkar*, *jugonogometaš*; *jugonogometen*, *jugošporten*), *kvazi-* (*kvazikarate*, *kvazišportnik*; *kvaziodbojgarski*, *kvazišporten*), *mini-* (*minikošarka*, *miniodbojka*, *minirokomet*, *minitenis*), *multi-* (*multišporten*), *neo-* (*neošportnik*), *para-* (*parašporten*), *psevdo-* (*psevdoatletika*; *psevdoplesen*), *turbo-* (*turboatlet*), *vele-* (*velenogometaš*; *velegolfski*, *velekolesarski*, *velenogometen*, *velesmučarski*, *velešporten*).

The following compounds belong to the category of the so-called suffixoid formations which have recognizable morphemes in the second part of the word:

*-kracija* (*športokracija*), *-log* (*nogometolog*, *športolog*), *-logija* (*športologija*), *-man/-men* (*golfsman*, *športmen*), *-manija* (*golfsmanija*, *nogomanija*, *nogometomanija*), *-oid* (*športoid*), *-slovec* (*nogometoslovec*, *športoslovec*).

Compounds in the category of compounds with combined interfixation and suffixation, which are mostly adjectival subordinate compounds, have a syntactic base which regularly contains a nominal phrase with a premodifying adjective (*ta, ki je povezan s smučarsko rekreacijo* (he or she who is connected with ski recreation) > *smučarsk-o-rekreacij-ski*, *alpinističnosmučarski*, *plesnobaleten*, *plesnodramaturški*, *smučarskotekaški*, *športnoaktiven*, *športnojahalen*, *športnokomentatorski*, *velikonogometen*, *vodnokošarkarski*, *zabavnošporten*).

The group of compounds with interfixation contains both coordinate (*sociološko-nogometološki*), and even more frequently, subordinate adjectival and nominal compounds (*boksaerobika*, *golfliga*, *všešportnik*; *všešporten*, *všešportnik*). The last example mentioned (*všešportnik*) could also be classified as a nominal derivation from an adjectival base.

### 5.1.3 Derivatives with prefixation

Due to globalization trends, there has been an increase in the number of hybrid derivatives with suffixation, i.e. derivatives which combine an international prefix with the native base, such as the following examples: *anti-* (*antinogomet*, *antinogometaš*, *antiples*, *antišportnik*; *antigolfovski*, *antinogometen*, *antišporten*), *post-* (*postples*; *postkolesarski*), *super-* (*supernogometaš*, *supersmučar*; *superatletski*), *ultra-* (*ultrakolesar*, *ultrašport*, *ultrašportnik*; *ultrakolesarski*, *ultrašporten*). Derivatives with native prefixes are also very productive: *nad-* (*nadatlet*, *nadnogometaš*; *nadkošarkarski*, *nadnogometen*), *naj-* (*najkošarkar*, *najnogometaš*), *ne-* (*nehokejist*, *nekegljač*, *nekolesar*, *nekošarkar*, *nekošarkaš*, *neples*, *nešport*, *nešportnik*; *nehokejski*), *pra-* (*prahokej*, *pranogomet*, *prasmučar*), *proti-* (*protikolesarski*, *protinogometen*), *so-* (*sokolesar*). In the category of verbal derivatives with prefixation (Vidovič Muha 1993, 2009), there is an interesting group of verbal derivatives with the following prefixes: *ob-* (*obkolesariti*), *po-* (*ponogometiti* < \**nogometiti*, *pokolesariti*, *posmučati*), *pri-* (*priodbojkirati* < \**odbojkirati*, *pribowlati*), *pre-* (*presankati*) in *za-* (*zasankati se*).

## 5.2 Potential complex words with a marked morphemic combination

**5.2.0** These complex words make use of predictable word-formational patterns, but they stand out due to their intentionally marked combination of bases and/or affixes. There are several kinds of markedness of combinations.

**5.2.1** In the field of sports, markedness in terms of time can be observed in the type of subordinate nominal compounds with interfixation and suffixation (*nogometoljubec*, *plesoljubec*, *športoljubec*), thus reviving the word-formational pattern which is in general becoming less and less productive. In SSKJ, for example, the compound *knjigoljubec* is labeled as obsolete; the compound *plesovodja* is similarly obsolete. SSKJ thus contains only 21 compounds of the type *knjigovodja*, *žerjavovodja*, etc.

**5.2.2** Sociolectal markedness can be reflected in the base or the affix of the word, or in both at the same time. Here we use the definition of sociolect as an established manner of communication within a specific group for the purpose of expressing its own specific perspective and value system (Skubic 2005: 173). Based on the corpus data from this study, we can distinguish at least two major groups of people creating the majority of such marked words, namely journalists, on the one hand, and sports enthusiasts on the other hand. The former use stylistically marked vocabulary in order to attract attention, while the latter use it in order to express their attitude toward a certain variety of sport. In order to determine where connotative lexis comes from and why it was created, the information about the sources in which such words appear can be very helpful. Such information can help to determine whether the studied material represents the word-formational core or the periphery of lexis. It is less probable to find stylistic lexis in technical and scientific literature, because its usage would clash with the main characteristics of expression in this kind of discourse. We therefore assume that in the case of the material from our study, stylistic neologisms belong to marginal lexis, which can be found in non-technical texts. A



more detailed corpus analysis has confirmed this assumption. The majority of new words from the field of sports come from magazines and newspapers, which FidaPLUS does not label specifically (only the label *magazine* or *newspaper* is used). The newspaper *Dnevnik*, its special editions *Nedeljski Dnevnik*, *Hopla* and *Pilot*, the newspapers *Delo*, *Večer*, *Vestnik* and *Štajerski tednik* are highly productive sources of new words, while a few examples can also be found in *Novi tednik* and *Dolenjski list*. Among magazines, the most productive source of neologisms and nonce words is the magazine *Joker*. This is not surprising, as *Joker* is a computer magazine whose Wikipedia entry states that it makes extensive use of archaisms and neologisms. According to Logar Berginc (2004: 196), the use of new complex words in this magazine is an important part of its image. New words also appear in the magazines *Mag*, *Golf Slovenija*, *MM marketing magazine*, *Premiera* and *Mariborčan*, and in some online magazines, although to a lesser degree. Taking this into account, the studied words can be characterized as so-called topical new words (“aktualistične nove besede”) (Bokal 2000), or journalistic ‘topicalisms’ (“poročevalski aktualizmi”) (Korošec 1998). It is interesting that the new vocabulary is created in equal measure by newspapers, which are characterized by natural and objective reporting about current topics, and by thematic magazines intended for specific interested audiences, which are usually freer in their expression. New vocabulary from the field of sports is to a large degree created by professional sports journalists (*Delo*, *Dnevnik*...), but sports enthusiasts and fans of a particular variety of sport are equally creative (*Joker*, *Golf Slovenia* ...).

Marked bases are found in complex words with parallel bases. Such doublet bases appear with *nogomet/fuzbal* (football/footie), yielding nominal derivations (*fuzbalček*, *fuzbalerščina*, *fuzbalščina*) from *fuzbal*, *fuzbaler* and *fuzbalski*. This parallel base also yields an adjectival derivative with prefixation (*antifuzbalski*), a prefixoid compound (*jugofuzbal*), and even a group of comopunds with a recognizable suffixoid (a final combining form) *-fen* (*fuzbalfen*), *-logija* (*fuzbalogija*), *-log* (*fuzbalolog*), *-manija* (*fuzbalomanija*). There is a similar case with *košarka/basket*, which yields the formation of the compound *minibasket* and hybrid derivatives (*basketiranje*, *basketiški*, *basketiš*). It is interesting to note that Bajec (1950: 112) observed that in the middle of the 20<sup>th</sup> century derivation with the suffix *-aš* was productive particularly in sports terminology (*kajakaš*, *nogometaš*, *odbojkaš*, *športaš*) under the influence of the neighboring Croatian language. A similar parallel word-formational pattern can be observed in *skijanje*, *skijač*, *skijaš*, *skijaški*.

Markedness in derivatives with suffixation and compounds with combined interfixation and suffixation can be observed in the case of alternative suffixes, with the base remaining unchanged: *-(ik/ež)*: *športnež*, *-(aš/nik)*: *nogometnik*, *-(aš/ec)*: *rokometec*, *-(ist/ec)*: *karatejec*, *-(ist/ač)*: *golfač*, *-(ist/i)at)*: *golfiat*, *-(lec/ač)*: *plesač*. Even a new prefixoid (initial combining form) *nogo-* is being created (*nogometafizika* < *metafizika nogomet*). A special and considerable group consists of nominal derivatives of neuter gender with the suffix *-nje*. Such derivatives are regularly deverbal in the system of Slovene word-formation, but in this case they are denominal, as it is impossible to extract a verbal base (*\*nogometariti* > *nogometarjenje* ‘playing foot-

ball over the internet'). In this case the potential base is the noun *\*nogometar* or *nogomet*. If we assume that the base is nominal, the suffix *-arjenje* can be interpreted as the result of perintegration.

A similar type of denominal derivation can be observed in *košarkašenje*, *tenisačenje*. With examples such as *\*košarkirati* (*košarkiranje*, *basketiranje*, *fitnesiranje*, *tenisiranje*), the implied bases are hypothetical verbs of the type *-irati*. The adjectival formation *nogometalen* is also stylistically marked, because it replaces the regular suffix *-en* with the deverbal suffix *-alen* which is used for denominal adjectival derivatives.

Markedness of the base and the affix can be observed in the example of compounds (*športoholičen*, *boksarskonazorski*, *plesolomen*, *tenisšpilen*) and in the feminine derivative with the suffix *-čarka* (*plesoholičarka*), for which the base *\*plesoholičar* or *\*plesoholik* has not been found in the corpus.

It is interesting to note that despite today's flood of the so-called e-formations, only one such example in low frequency lexis from the field of sports could be found (*e-smučarščina*).

### 5.3 Nonce words

**5.3.1** Nonce words are characterized by the use of individual word-formational patterns and considerable creativity, also through the use of word play. Such examples manipulate existing syntactic structures, but use morphological operations which do not exist in natural language (Bat-El 2000: 66). The ad-hoc nature of such formations is shown in the fact that they do not fill any lexical gaps, because there already is a stylistically neutral expression in the lexical system. The established expression, however, cannot carry over the author's conscious expressiveness.

**5.3.1** Word play arises from the phonemic similarity of two words, e.g. *golfiga* (*Golf je navadna gol(ju)fija*), *golfivec* (*Nikoli ni igral golfa, sedaj pa igra kot gol(ju)fivec*), *golfissimo* (*Neverjetno! Bravo, bravissimo! To je golfissimo!*). At its base there can also be a word phrase which is truncated; the juxtaposed truncated elements preserve the overlapping element, which is similar to an interfix. Here we could use the term blends ("prekrivanke"), which are similar to compounds in that they consist of at least two lexemes but are formed according to a pattern which contains more irregularity than the pattern for creating compounds (Dressler and Merlini Barbesi 1994, quoted in Bat-El 2000: 63). Algeo (1977: 48, quoted in Lehrer 2007: 1116) defines blends as compounds which consist of one word and a part of another word or two (occasionally three) parts of words or truncations which cannot exist as independent units (Lehrer 2007: 116), e.g. *nogogazem* (< *nogo(metni or)gazem*), *smučarajanje* (*smučar(sko) rajanje*), *smučanorija* (*smuča(rska) norija*), *smučerator* (< *smučar(ski koment)ator*), with the replacement of *a* for *e*.

**5.3.2** Many nonce words are formed as a consequence of the use of a new motivating word, for example in the case of football and its word family (*nogobrc*, *nogobrcar*, *nogohitrec*, *nogoljublje* 'love of football', *nogovanje* 'artistic playing of football') or ball sports

in general (*žogošport, žogošporten*). In similar cases there is a replacement of a foreign morpheme by a native morpheme (*brcoboks < kimboks, brcoboksar < kimboksar*). The flouting of syntactic rules can be observed in the formation of gerunds, which would normally be formed from imperfective verbs, as in the example *prekolesarjenje*. The corpus also shows no examples of the verb *\*razkolesariti* for the past participle *razkolesarjen*.

There are also examples of atypical word-formational analogy, such as *slačinogometaš* on the analogy with *slačipunca*, *golfizem* on the analogy with *turizem* (e.g. *Madeira je in bo v golfu ena izmed mek evropskega zimskega golfizma, čeprav je res, da je cenovno v precej višjem razredu v primerjavi s celinskimi igrišči v Španiji, na Portugalskem, da o cenenih golfskih počitnicah v Tuniziji in Turčiji ne govorimo.*). There are also atypical combinations of root constituents in compounds with the zero interfix (*golfhec, golfsafari, rolerhokej, karateseks*), in addition to those with the more usual interfix *-o-* (*nogoodbojka*).

**5.3.3** The corpus data also contains the multi-constituent coordinate adjectival compound *rokometnonogometnekošarkarskokolesarski* (*Veseli nas, da ste tudi član rokometnonogometnekošarkarskokolesarske rekreativne ekipe*), which is spelled as a juxtaposed compound, but can only be interpreted as (a member of) a *handball-football* and *basketball-cycling* (recreational teams), and would normally have to be spelled using a blank space and two hyphens.

## 6. CONCLUSION

The study of new complex words in the field of sports presented in this article focuses on the structural-semantic analysis. In terms of word-formation, these words mostly belong to the category of potential complex words, while there are fewer actual nonce words, i.e. individual complex words which are formed according to original patterns. In the category of potential complex words, prefixoid compounds (*kvazišportnik, psevdoatletika*) and suffixoid compounds (*nogometolog, golfmanija*) are predominant. Nonce words are strongly stylistically marked, because in most cases they exist as parallel expressions for denoting phenomena which already have a name. (e.g. *brcoboks = kimboks, nogobrca = nogometaš, tenisarka = tenisačica*). This has also been confirmed by the analysis of the sources in which such new words appear. Most of the examples of such new words have been found in daily newspapers and magazines, while only a few examples have been found in other sources, such as fiction (e.g. *nogometaška, plesač*) or in literature labeled in the corpus as non-fiction literature (e.g. *sankaštvo, smučljiv, tenisirati*). Authors use word-formational creativity to make their texts more attractive and to have more influence on the readers, who are expected to recognize the stylistic and new meaning of such words. This cannot always be observed solely by focusing on word-formational constituents, so the situational and textual context also has to be taken into consideration. So new complex words, in addition to their expressive function, also have an intellectual function, and both are expressed either by sports enthu-

siasts who show their love of sports, or by professional journalists who strive for originality in their writing. The professed general purpose of all journalistic news reporting is to inform publicly and objectively about current events, but in today's world of fierce competition even serious news reporting has been influenced by commercialization and even tabloid culture. As a consequence, journalists, similar to advertisers, try to attract attention through linguistic creativity. The following two conclusions can thus be reached: (1) journalistic texts remain the most important source of more recent or unusual words also in the case of sports; (2) the occurrence of neologisms in journalistic texts, which contrary to technical and scientific texts use affective and expressive style to a higher degree, shows that the analyzed examples mainly belong to the peripheral areas of the lexicon, broadening Slovene sports lexis either through not yet established and marked morphemic combinations or through new word-formational patterns.

### References:

- BAJEC, Anton (1950) *Besedotvorje slovenskega jezika I: Izpeljava samostalnikov*. Ljubljana: Ljudska pravica, SAZU.
- BAT-EL, Outi (2000) "The grammaticality of 'extragrammatical' morphology." In: U. Doleschal, A. M. Thornton (eds), 61–84.
- BOKAL, Ljudmila (1998) "Tipologija novih besed (ob primeru avtomobilskega izrazja)." In: M. Humar (ed), 147–162.
- BOKAL, Ljudmila (2000) "Aktualistične nove besede v slovenščini." In: Z. Jan (ed), 68–76.
- FICKO, Simona (2009) *Neologismen im Bereich Sport (ein deutsch-slowenischer Vergleich)*, Diplomarbeit. Maribor: Filozofska fakulteta.
- KALIN GOLOB, Monika (2008) "SMS-sporočila treh generacij." In: M. Košuta (ed), 283–294.
- KOROŠEC, Tomo (1998) *Stilistika slovenskega poročevalstva*. Ljubljana: Kmečki glas.
- KRIŽAJ ORTAR, Martina/BEŠTER TURK, Marja et al. (2000) *Na pragu besedila 2. Učbenik za slovenski jezik v 2. letniku gimnazij, strokovnih in tehniških šol*. Ljubljana: Rokus.
- LEHRER, Adrienne (2007) "Blendalicious." In: J. Munat (ed), 115–133.
- LOGAR BERGINC, Nataša (2004a) "Nove tehnologije in nekateri nesistemske besedotvorni postopki." In: E. Kržišnik (ed), 121–132.
- LOGAR BERGINC, Nataša (2004b) "Nove tvorjenke v publicistiki." In: M. Kalin Golob, M. Poler Kovačič (eds), 175–199.
- MARTINCOVÁ, Olga (1972) "K problematice lexikalnih neologizmu." *Slovo a slovesnost* 33/4, 283–293.
- MUHVIĆ-DIMANOVSKI, Vesna (2005) *Neologizmi. Problemi teorije i primjene*. Zagreb: Zavod za lingvistiku Filozofskega fakulteta Sveučilišta u Zagrebu.
- NIKOLIĆ-HOYT, Anja (2004) *Konceptualna leksikografija: prema tezaurusu hrvatskoga jezika*. Zagreb: Hrvatska sveučilišna naklada.
- Revija Joker*. 18 June 2011. [http://sl.wikipedia.org/wiki/Joker\\_\(revija\)](http://sl.wikipedia.org/wiki/Joker_(revija)).
- SKUBIC, Andrej E. (2005) *Obrazi jezika*. Ljubljana: Študentska založba.

- STRAMLJIČ BREZNIK, Irena (2003) "Besedotvorna tipologija novonastalega besedja s področja mobilne telefonije." *Slavistična revija. Zbornik referatov za XIII. mednarodni slavistični kongres, 15.–21. avgust 2003*, 105–118.
- STRAMLJIČ BREZNIK, Irena (2004) *Besednodružinski slovar slovenskega jezika. Poskusni zvezek za iztočnice na B*. Maribor: Slavistično društvo Maribor.
- STRAMLJIČ BREZNIK, Irena (2007) "Historiat poimenovalne in pojmovne raznolikosti za leksemske združbe v slovenistiki". *Časopis za zgodovino in narodopisje* 78 = n. v. 43/1, 97–109.
- STRAMLJIČ BREZNIK, Irena (2010) "Besedotvorne lastnosti slovenskih okazionalizmov." In: E. Vasil'evna Petruhina (ed), 352–363.
- STRAMLJIČ BREZNIK, Irena/VORŠIČ, Ines (2009) "Grafoderivati v tiskanih oglasih." *Teorija in praksa* 46/6, 826–838.
- TOPORIŠIČ, Jože (1992) *Enciklopedija slovenskega jezika*. Ljubljana: CZ.
- TOPORIŠIČ, Jože (1996) *Slovenski jezik in sporočanje 2*. Maribor: Založba Obzorja Maribor.
- TOPORIŠIČ, Jože (2000) *Slovenska slovnica*. Maribor: Založba Obzorja.
- TRIER, Jost (1931) *Der deutsche Wortschatz im Sinnezirk der Verstandes: Die Geschichte eines sprachlichen Feldes, Band I: Von den Anfängen bis zum Beginn des 13 Jh.* Heidelberg.
- VIDOVIČ MUHA, Ada (1993) "Glagolske sestavljenke – njihova skladenjska podstava in vezljivostne lastnosti: z normativnim slovensko-nemškim vidikom." *Slavistična revija* 41/1, 161–192.
- VIDOVIČ MUHA, Ada (2000) *Slovensko leksikalno pomenoslovje: govorica slovarja*. Ljubljana: Znanstveni inštitut Filozofske fakultete.
- VIDOVIČ MUHA, Ada (2009) "Skladenjska interpretacija glagolskih predponskih obrazil – vprašanje propozicije." *Slavistična revija* 57/2, 251–261.
- VUKOVIĆ, Petar (2011) "Haller, Jakobson i jezična globalizacija *ante litteram*." In: B. Kryžan-Stanojević (ed), 19–38.

### **Linguistic sources:**

- DORNSEIFF, Franz (2004) *Der deutsche Wortschatz nach Sachgruppen*. Berlin: W. de Gruyter.
- FidaPLUS*. 10 June 2011. <http://the.sketchengine.co.uk/>.
- Seznam športov*. 15 June 2011. [http://sl.wikipedia.org/wiki/Seznam\\_%C5%A1portov](http://sl.wikipedia.org/wiki/Seznam_%C5%A1portov).
- Sketch Engine*. 10. June 2011. <http://the.sketchengine.co.uk/>.
- SSKJ = *Slovar slovenskega knjižnega jezika z Odzadnjim slovarjem slovenskega jezika in Besediščem slovenskega jezika z oblikoslovnimi podatki*, Elektronska izdaja na plošči CD-ROM. Ljubljana 1998.
- SP = *Slovenski pravopis*, 2001. J. TOPORIŠIČ (ed). Ljubljana: SAZU in Znanstvenoraziskovalni center SAZU.

Abstract  
WORD-FORMATIONAL PRODUCTIVITY OF THE SLOVENE LANGUAGE  
IN THE CASE OF SPORTS NEOLOGISMS

Newly created words are a frequent topic of research in modern linguistics. The reason for this is that innovation processes in lexicon are a reflection of dynamic tendencies which occur as a consequence of modern communication requirements. This contribution focuses on the field of sports, using a corpus approach based on the Sketch Engine (SkE). It investigates formations in Slovene with low frequency (e.g. *nogometoljubec*, *karatejka*, *sankijada*). Based on their morphemic structure, it determines whether they are potential neologisms, which are created according to a productive pattern, or nonce formations, which are created according to individual patterns connected with specific text usage, and which as ad-hoc formations do not belong to the productive word-formational patterns. In light of these findings, the question of the combinations of domestic and foreign word-formational elements will be discussed, and consequently also the question of the hybridization of Slovene formations today. The corpus-based approach will enable us to determine whether the analysed material represents a word-formational core or periphery of the lexicon.

In order to detect new lexical processes, it is necessary to investigate atypical word-formational processes which open up the creative potential of language. They testify to the dynamic nature of language and to its capability to respond to the different needs of different users in different circumstances.

Povzetek  
BESEDOTVORNA PRODUKTIVNOST SLOVENŠČINE  
NA PRIMERU NOVOTVORJENK S PODROČJA ŠPORTA

Raziskovanje novonastajajočih besed predstavlja aktualno področje sodobnega jezikoslovja, saj so inovacijski procesi v leksiki odsev dinamičnih teženj, ki jih narekujejo sodobne komunikacijske potrebe. Prispevek na primeru pomenskega polja *šport* s korpusnim pristopom v orodju SkE prikazuje tvorjenke z nizko frekvenco (npr. *nogometoljubec*, *karatejka*, *sankijada*) in na podlagi njihove morfemske strukture ugotavlja, ali gre za potencialne novotvorjenke, nastale po produktivnih vzorcih, ali za priložnostnice, tvorjene po individualnih vzorcih, vezanih na specifično besedilno rabo, zato pa poimenovalno fakultativne ter brez sistemotvorne možnosti. S tem bo nakazano vprašanje kombinatorike domačih in prevzetih tvorbenih prvin in vprašanje hibridizacije slovenskih tvorjenk danes. Uporabljeni korpusni pristop bo z analizo virov omogočal diagnosticirati, ali gradivo predstavlja besedotvornovzorčni jedrni ali periferni del leksike.

Če hočemo detektirati nove leksikalne procese, se mora besedotvorje nujno zanimati za jezikovnosistemsko besedotvorno atipiko, ki v jezik prinaša novi tvorbeni nemir in ustvarjalni potencial, s katerim se dokazuje živost jezika ter njegova sposobnost odzivanja na različne potrebe različnih uporabnikov v različnih okoliščinah.

## BEYOND DICHOTOMIES: ON THE NATURE AND CLASSIFICATION OF COMPOUND VERBS IN ENGLISH

### 1. ON CONSTRUAL

Language as a human cognitive capacity is one of the basic tools for externalizing the way we conceptualize the world for communicative purposes. No matter how complex and encyclopedically rich our situated knowledge is, in “thinking for speaking” (Slobin 1987: 435; 2003: 158) we constrain our knowledge in ways that the symbolic inventory of our language affords to us. This symbolic inventory reflects the numerous attentional acts that have been conventionalized in alternative construals (Evans 2009, Langacker 2008, Marchetti 2006, Talmy 2000).

*Construal* captures the conceptual-linguistic interface and describes how elements from the symbolic inventory of language engender specific manipulations of cognitive content by the cognizing subject. Compound verbs (CVs)<sup>1</sup> represent special construal mechanisms<sup>2</sup> and their cognitive function is best understood if they are studied as actualizing a constructional idiom and constituting a cognitively homogeneous class. For the purposes of the current discussion by CV is understood a compound which irrespective of its derivational pattern (composition proper, conversion or back-formation) has as its core meaning a verbal meaning and can be used in all finite forms.

“[S]emantic analyses prove to depend crucially on what *construals* of objects and events are possible and salient” (Jackendoff 2010: 8; emphasis added). The structuring of linguistic concepts is contingent on the distinction between event construals and object construals (Evans and Green 2006; Langacker 1987, 2008; Radden and Dirven 2007, Talmy 2000), which hinges on alternative acts of profiling executed by the activation of features of several construal types, with only the ones most relevant for the study of CVs listed below:

---

\* *Author's address*: Faculty of Classical and Modern Philology, Department of British and American Studies, 15 Tsar Osvoboditel Blvd., 1504 Sofia, Bulgaria. Email: abagasheva@gmail.com

<sup>1</sup> The argument in the paper is based on the lexico-semantic analysis of a self-compiled corpus of 239 CVs. All primary sources from which the CVs have been excerpted are listed as primary sources in the References section.

<sup>2</sup> CVs belong to the *composites* class of linguistic units as defined by Lampert (2009: 62-63) “[t]wo categories relevant for linguistic representations at all levels [...] must [...] be kept apart: First, those that result from an ‘additive’ (or: computational) combination of semantically and/or formally simplex items, yielding, [...], *compositions* of variable complexities in accordance with combinatorial rules; second, there are *composites*, which cannot readily be analyzed in terms of a ‘simple’ (additive) computation of their formal constituents and/or semantic components, but only as ‘wholes’ or Gestalts.”

Figure-ground<sup>3</sup> (profile-base opposition)  
Granularity – filtering against a frame  
Degree of schematicity

“[C]onstrual phenomena [...] variously impose structure on conceptualizations in ways that do not immediately follow from their content” (Verhagen 2007: 51). Symbolic units mould conceptualizations and impose focal adjustments by instructing interactants as to which are the central components to be attended to in a given conceptualization accessed via that particular symbolic unit. In analyzing different types of construal phenomena two dimensions of particular pertinence<sup>4</sup> for the study of CVs are *frame of knowledge* (or ICM) and *composition* of the conceptualization. Verbs are by definition focal adjusters as their role “in a construction is to provide specifics to the schematic conceptualization evoked by the construction” (Verhagen 2007: 51). CVs are special in terms of *composition*, since they combine a thing/property and event construal for which they provide special constructional scaffolding. This specificity determines the marked degree of granularity associated with CVs which expand the profiling of a frame and which, in terms of figure-ground perspectivization, might (or at least one of the CV prototypes in English does) include as foregrounded a chosen component of an already profiled event (see 4.3. and 5).

CVs in English provide a finely-granulated structure to conceptualizations which either specifies values for variables from underlying frames (subclassification CVs) or create novel blended conceptual spaces (superclassification CVs), i.e. novel conceptualizations. A heightened resolution effect can only be observed with subclassification verbs (e.g. *drip-dry*, *deep-fry*, *spoon-feed*, etc), which are characterized by lowered schematicity as they are bi-focal perspectivizers. Superclassification verbs on the other hand are not characterized by reduction in schematicity as their semantics captures a novel conceptualization (e.g. *brown-bag*, *hag-ride*, *moon-light*, *high-tail*, etc.). The two types constitute construal CV prototypes, which with their contingent radial networks instantiate the constructional idiom CV.

Verhagen (2007: 48) contends that “the best way to make these notions [of construal phenomena] relevant for linguistic analysis is not given a priori and thus requires empirical investigation.” In keeping with such desiderata in the remainder of the paper CVs are analyzed with the aim of elucidating their special properties as

---

<sup>3</sup> It is of paramount importance in the study of CVs to distinguish between concept-type profiling (i.e. external) and compound internal figure-ground profiling. The first kind of profiling relates to the profiling of the verbal concept against the frame which it is evoked by and which it evokes, while the second captures the constituency-internal relative profiling within the emancipated profile (i.e. composition).

<sup>4</sup> These are the two parameters generally utilized for the initial classification of different kinds of construal phenomena. Among the focal adjustments suggested by Langacker (1987: 116-37), Talmy (2007: 276-277) and Croft and Cruse (2004: 46; 63-66) extremely pertinent for the analysis of CVs are components of the perspective construal variables – *figure/ground*, *configuration (composition)* and *constitution/gestalt* (or *profiling*).



construal mechanisms used to conceptualize *situations* or rather *conceptual cores* as finely granulated *relations*.

In discussing the ways in which “units of thought relate to language” Radden and Dirven (2007: 41) claim that we do not need “more than two basic types of conceptual units things and relations” in order to establish linguistically relevant conceptual distinctions. These are combined into complex “conceptual cores” that are ultimately expanded to “situations” (i.e. clausal descriptions). In view of this claim CVs pose a problem for neat dichotomous treatment as they very often both conceptually and in terms of form include a “thing” (e.g. *flat-hunt*, *house-sit*, *fellow-feel*, *case-harden*) and thus come closer to a “conceptual core”/“situation” than to a “relation”. It is exactly because of the fact that CVs profile/perspectivize “situations” as “relations” that they appear to possess both morphological and syntactic properties and to structure in a “relational” profiling conceptual content partaking of a “situation”.

## 2. ON COMPOUND VERBS

Masini (2009: 256) holds it that

compounds may be the product of both productive rules and lexicalization processes, whereas phrases are fully productive; also, compounds may be both semantically transparent and idiomatic, whereas phrases are generally transparent; finally, parts of compounds cannot undergo syntactic operations.

Compounds seem to differ significantly from phrases. At the same time it has been stated that they have internal structure which displays syntactic properties<sup>5</sup>, i.e. “genuinely syntactic combination of lexical items” (Anderson 1992: 292), so they can’t possibly also be simply words. Thus CVs appear to realize constructional idioms both in terms of construal mechanisms and in terms of their position within the hierarchical network of the symbolic inventory.

Schmid (2011: 122) contends that “there are no typical verbal compounds in English” in terms of the following criteria:

- i) derivational morphology – bi-constituency;
- ii) inflectional morphology – grammatical categories marked on the head constituent;
- iii) phonology – single intonation unit, with main stress on the modifier;
- iv) semantic structure – not strictly compositional and semantically lexicalized;
- v) sociopragmatic perspective – naming, not deictic function, high level of institutionalization;
- vi) cognitive perspective – entrenched concept, stored in the mental lexicon (ibid. 121-122)

---

<sup>5</sup> “Compounding ... involves the combining of stems from the lexicon into a *quasi-syntactic structure*. This word-internal structure seems to be *unique* to compounds, in fact.... The formation of compounds seems to involve a *genuinely syntactic combination* of lexical items below the level of the word (perhaps along lines like those explored in Selkirk 1982), while non-compounds have only a phonological structure.” (Anderson 1992:292; emphasis added).

N + V	carbon-copy, carbon-date, babysit, head-hunt, blockbuster
A + V	free-associate, double-book, soft-land, crash-land, fine-tune, whitewash
Prt + V	outnumber, overachieve, overeducate, overcook, underrate
V+V	drink-drive, crash-land, dry-clean, stir-fry
A + N	brown-bag, bad-mouth, blacklist, mainstream
N + N	breath-test
Num+V	double-cross, double-check

Actually, CVs are not typical compounds according to one of the criteria in the above list<sup>6</sup>. CVs might be considered atypical compounds in relation to derivational morphology as they constitute a heterogeneous class resulting from conversion, back-formation and compounding, even though they conform to the bi-constituent criterion. In view of this conflict, it might do justice to CVs if we make provisions for accounting for compound lexemes and compounding as a word-formation pattern separately. The grounds for such a step can be found in the following:

- a) the uniformity of the nature of the meaning composition in diverse CVs (where diversity is captured in i) internal constituency (NV; AdjV, VV, PV, NN(v), NumV); ii) pattern of derivation (conversion, back-formation, analogy, compounding); iii) development of the word-formation paradigm associated with the CV (agent noun, activity noun, etc.);
- b) distinctions between the cognitive function of a word-formation pattern and of types of complex words (Schmid 2011: 18-19);
- c) a distinction in degree of ecological validity can be postulated between a word-formation pattern (compounding) and its products (compound lexemes). In the f-mind of speakers the latter are a lot more prominent, directly available and amenable to various cognitive operations, while the former is more likely to be part of a logical and neat linguistic model which accounts for the meaning of complex lexical items in language;
- d) after all “the goal is an account of compound meaning that is as rich as the account of word meaning” Jackendoff (2009: 115).

With the advent of cognitive linguistics, it has been established that conversion mechanisms operate at the conceptual level and surface as profiling preferences in construal operations (Evans and Green 2006; Langacker 1987, 2008). Thus conceptualizing and linguistically encoding a knowledge structure as a nominal or verbal lexeme is an underspecified option of alternation (Farrell 2001). Admitting the possibility for alternative encodings does away with the necessity to account for the cog-

---

<sup>6</sup> The phonological criterion is not taken into consideration as it constitutes the form pole in the symbolic pairing of meaning and form and is considered beyond the scope of the present argument.

nitive effects of a specific word-formation pattern and allows us to concentrate on the products, i.e. on CVs as construal mechanisms.

The special cognitive function of CVs thesis can easily be corroborated on semantic grounds. Cruse (2004: 222-227) in defining philonyms, xenonyms and tautonyms, postulates that “it is the duty of a non-head to bring information not available in the head.” If the semantic non-head brings into the composition information which is already contained in the head, the result will be a tautonymy. In some cases it is possible for a selectee of the semantic verbal head (selector) to express meaning already contained in the head, on condition that the selectee highlights a special aspect of the action denoted by the verb (e.g. shrug one’s shoulders vs. \*smile one’s lips).

Many CVs contain as first constituent a foregrounded element of the frame which an associated simplex verb profiles (e.g. *deep-fry*, *force-feed*, *sun-dry*, etc.). Contrary to expectations for pleonastic semantic effects, even such subclassification CVs are meaningful and informative, which requires that the foregrounded element be relevant, unexpected, unpredictable or *highly specific*. Kiefer’s (1993: 46) definition of argument relevance to the head in a compound best generalizes these requirements:

An argument in a compound is said to be relevant with respect to the head if it is not predictable on the basis of the meaning of the head and world knowledge. It is also possible to define a scale of relevance: the more predictable an argument is with respect to a head the less relevant it is.

This broad definition dependent on world knowledge is further refined by the postulation of a “a scale of relevance on the basis of the range of possible arguments. The wider the range R of the arguments  $A_i$  admitted by the head H is, the more relevant an argument  $A_k$  out of R will be” (Kiefer 1993: 50). Relevance is associated with and motivated by ontological salience<sup>7</sup>.

In this way the range of foregrounded constituents in a CV appears to be regulated by system-external constraints, most probably stemming from general cognitive abilities. As Kiefer (1993: 55) himself admits, “the selection of arguments in compounds is thus only in part a matter of grammar, it is to a large extent determined by extralinguistic considerations.” Besides being recognized as heavily influencing compounds, system-external considerations have been pointed out as an indispensable analytical heuristic by Bundgaard et al. (2006: 369) who insist that “any attempt to define a combinatorial rule in terms stemming exclusively from the linguistic system as such (*qua* a self-contained formal system) is doomed to fail.”

---

<sup>7</sup> “[O]ntological salience,” is not related to temporary activation states of concepts but to more or less stable properties of entities in the world. The idea is that by virtue of their very nature, some entities are better qualified to attract our attention than others and are thus more salient in this sense” (Schmid 2007: 120).

### 3. FRAMES, CONCEPTUAL CORES AND CVS

The most widely accepted operationalization of extralinguistic factors that have direct bearing on linguistic units is the *frame* as the most prominent unit of conventionalized background knowledge at the conceptual level. Within the cognitive linguistic enterprise it has become customary to assume that “[f]rames structure our conceptual and social life” (Fauconnier and Turner 1998: 236). Frames underlie the profiling and resultant construal in different pairings of meaning and form. Fillmore (2006: 378) defines the correlation between construal mechanisms and lexical items as a mutually implicating one,

[u]sing the word ‘frame’ for the structured way in which the scene is presented or remembered, we can say that the frame structures the word-meanings, and that the word ‘evokes’ the frame.

CVs profile frames in two distinct ways, either by constituent foregrounding (subclassification CVs) or by engendering a novel blend (superclassification CVs). The type of frame which is captured in a CV is the *event schema*. The latter is defined as the abstracted, recurrent conceptual pattern (prototype) we associate with situations. As Dirven and Verspoor (2004: 75) state “our language shows that we tend to group events according to a limited number of types, called “event schemas””. These are traditionally encoded in clausal patterns in which arrangement of symbolic units is aligned with anthropocentricity-driven salience profiling. Simplex verbs profile a temporal relational concept exclusively, while CVs profile a larger part of the “conceptual core” of a situation. Since “[p]rofilng amounts to nothing more than the relative prominence of substructures within a conceptualization, and is inherently a matter of degree” (Langacker 1990: 208), CVs seem to occupy a fuzzy area between relational concepts and conceptual cores, i.e. they constitute construal mechanisms unto themselves.

The best way to illustrate<sup>8</sup> this claim is to compare the profiling provided by a simplex verb and a compound verb both evoking the same frame. The degree understanding of profiling allows for the postulation of a cline which captures the fine-grained degrees in which different CVs display “relational” vs. “conceptual core/situational” profiling.

Feature foregrounding:

[ss v]v

[T]he two old celebrants argued vociferously and at length about who should hold the knife to make the first cut.

[They] even out-argued or outmaneuvered that wild-eyed White House bunch for a change...

---

<sup>8</sup> All the sentence examples analyzed immediately below have been excerpted from Corpus of Contemporary American.

In the first sentence the verb names an activity of angrily showing one's disagreement with another participant in the situation. In the second, the embedding of the verb's frame in a spatial specifier frame in the constitution of the CV projects the feature TO DEFEAT OR GET THE BETTER OF by a complex metaphonymic scenario (described in 5).

[X,Y, Z V]V spoon-feed

*He fed himself on plankton, seaweed, and an occasional fish.*

*My friend Kristin spoon-fed her older son, Cole, until he was 4.*

In this non-metaphoric constitution of the CV, the feature-foregrounding is achieved by the emancipation of a secondary figure (the instrument), which accounts for the bi-focal perspectivization in the construal of the single-scope CV blend.

*Reality television has fed what he calls people's voyeuristic impulses.*

*In the new arrangement, reporters had to wait to be spoon-fed, and the White House made only minimal efforts to do even that.*

In the metaphoric constitution of the CV, the inclusion of the instrument in the construal creates a novel meaning feature relating to manner. The second sentence exemplifies the conceptual metaphor KNOWLEDGE IS FOOD, while the first one is based on INTEREST IS APPETITE, both rooted in the same source domain FOOD. The meaning of the CV is fully lexicalized as the IN A SIMPLIFIED MANNER semantic component cannot be directly computed by the combination of the constituents or through the conceptual metaphor itself.

Novel conceptualization:

[XY]V

*He footed it on dusty roads, or sped magnificently in freight cars, counting time as of no account.*

*So I doubt you need to pussyfoot around the issue as much as you think.*

It is difficult for a relation (even a metaphonymic one) to be established between the two sentences in the pair. The CV's meaning AVOID MAKING A DECISION OR EXPRESSING AN OPINION BECAUSE OF UNCERTAINTY OR FEAR is an entirely novel conceptualization which is not associated with any of the involved input frames separately. It arises out of a complex series of metaphonymies engendered in the elaboration of the blend.

As is obvious from the above analysis, CVs are onomasiologically special in relation to simplex verbs. They *name* a situation with salient features explicitly designated in the naming unit.

*Words are tools to pilot attention.* The elements composing *the meanings of words are attentional operations: each word conveys the condensed instructions on the attentional operations one has to perform* if one wants to consciously experience what is expressed through and by it (Marchetti 2006: 163, emphasis added).

The special status of CVs becomes extremely important, since in a clausal representation of a situation the attention is distributed over several pivots in the symbolic unit, while in a CV the feature/participant secondary trajector is condensed into a naming unit.

**4. THE SEMANTICS AND CLASSIFICATION OF CVS IN ENGLISH**

In such context<sup>9</sup> CVs are expected to have special lexical semantics due to their status as complex words involving composition. Actually, non-compositionality, or not strict compositionality, is promoted as a defining feature of compoundhood. Jespersen notes that “we may perhaps say that we have a compound if the meaning of the whole cannot be logically deduced from the meaning of the elements separately” (Jespersen 1954: VI, § 8.1.3). For the study of the whole we can adopt two distinct approaches – account for the semantics on the basis of the word-formation pattern recognized and start the analysis from the source or constituent elements or take the complex word as a unified symbolic unit and study its profiling and the lexical concept which results from that.

**4.1 Meaning mechanisms in conversion and back-derivation**

If we adopt the former approach, we need to postulate three different mechanisms of meaning generation for the three groups of CVs we have to recognize on the basis of different types of word-formation:

<i>CVs arising from</i>		
<i>back-derivation</i>		
babysit	←	babysitter/baby-sitting
bottle-feed	←	bottle-feeding
stage-manage	←	stage-managing
 <i>conversion</i>		
sandbag	→	sandbag
blackball	→	blackball
railroad	→	railroad

---

<sup>9</sup> Lexical meaning is a special type of data structure which is associated with an emergent or ad hoc concept for whose generation the lexical item functions as an instructional sign-post which directs the listener in constructing the ad hoc concept on the basis of the perceptually presented stimulus (the lexical form) with its prototypical lexical concept which might or might not undergo significant modification in the particular instance of language use of the units in the f-mind as tools to pilot attention (Evans 2006, 2007, 2009; Marchetti 2006; Jackendoff 2002).

*compounding/composition*

drip-dry  
small-talk  
fast-talk

Lieber (1981: 186) defines the following semantic interpretation rule for denominal, conversion verbs, “given a semantically specified noun X, and a related, but semantically underspecified verb Y, X must serve as an argument in the interpretation of Y”. Though this may work for simplex verbs, it often fails with CVs. It would be difficult for any linguistic theory to establish *railroad<sub>(n)</sub>* as an argument of the CV *railroad* with its meanings (OED): orig. U.S. TO ACCOMPLISH (AN ACTION) WITH GREAT SPEED; TO ‘RUSH’ (A PERSON OR THING) TO OR INTO A PLACE, THROUGH A PROCESS; TO HUSTLE, TO COERCE; TO SEND (SOMEONE) TO A PLACE OF PUNISHMENT WITH SUMMARY SPEED OR BY MEANS OF FALSE EVIDENCE.

The belief that the parent noun is necessarily involved in the meaning generation mechanism of a denominal CV leads to implausible lexical semantic interpretations of attested CVs: e.g. Nagano<sup>10</sup> (2007: 65) classifies as LOCATUM verbs *air-condition*, *face-lift*, *ill-treat*, *ill-use*, *pressure-treat*, *triple-tongue*, *turbocharge* and *valet-park*, while *tailor-make*, *jam-pack* and *hard-boil* appear as GOAL and *prize-fight* is classified as SOUND-SYMBOLISM. As is intuitively obvious such classification based on supposedly uniform meaning associated with word formation specifics do not actually reveal much about the semantics of the CVs.

The more or less agreed upon meaning generation mechanism for back-formed CVs is dependent on the fact that back-formation in compounding phenomena operates on synthetic compounds. The cognitive procedure involved is counterintuitive as back-formation for CVs occurs with secondary compounds in whose analysis linguists of quite diverse persuasions recognize a motivating link usually associated with argument readings of the first constituent, i.e. a verbal element is necessarily recognized in the initial composition of the synthetic compound. We can safely claim that a compound CV is postulated (though generally not acknowledged) in backstage cognition which drives the semantic computation/analysis of secondary compound nouns (both *-er* agentive and *-ing* activity ones) or adjectives. In their it is generally assumed that they,

do express a straightforward semantic relation. Secondary compounds are characterized by an argumental relation between the constituents: it is a logical condition of this type

---

<sup>10</sup> In the article “Marchand’s analysis of back-formation revisited: back-formation as a type of conversion” Nagano develops a theory of back-formation as conversion and adopts for the semantic classification of CVs Clark and Clark’s (1979) list for converted denominal verbs. Even though the theory does away with the necessity for distinct approaches to back-formed and converted CVs, it leads to unconvincing semantic interpretations of attested CVs. What is more, conversion is traditionally associated with root compounds giving rise to CV, while the bases in back-formation are classified as synthetic compounds.

of compound that at least one of the constituents is of verbal nature (i.e. a *pure V*, or a deverbal derivative)” (Guevara and Scalise 2004: 8).

In the composition of synthetic compounds a verbal component is necessary for the conceptualization of the cognitive content which is subsequently profiled or coerced as a nominal, thing-construal. It appears that in back-formation we exert cognitive effort to re-profile or reverse a profiling procedure from nominal construal back to an earlier stage in the meaning composition of the profiling. Thus the adoption of the word-formation pattern based approach leads to an impasse of unsatisfactory semantic analysis, which reveals the theory-dependent classificatory schemas used in categorizing compounds.

A way to avoid duplication of cognitive effort and analytical difficulties is to side with Farrell’s (2001) contention that nominal/verbal construal is a matter of alternative profiling of underspecified symbolic units which are related via functional shifts.

The lexical semantic representations of such words include event schemas that are compatible with either noun or verb meanings. The verb vs. noun aspect of the meanings is supplied by the morphosyntactic contexts in which they appear (Farrell 2001: 109)

## 4.2 Compound-internal constituents

Many of the controversies surrounding CVs stem from the tacit but deeply rooted assumption that the constituents of compounds are categorically-specified words or roots. We believe that such an assumption can be discarded for the sake of a unified analytical approach to the semantics and classification of CVs in English. Just as “[c]onceptual structure encodes our knowledge of the categorical status of auditory, visual, olfactory, and haptic information in a neutral or a-modal format” (Van Der Zee and Nikanne 2000: 4), language as one of the cognitive systems in the f-mind provides parallel options for reconfiguring conceptual content in cases of unifying schemata in compound phenomena. Further support for the actegorial status of compound internal constituents can be found in the understanding of part-of-speech assignment as pragmatically driven and constrained.

Hopper and Thompson (2004) put forward the hypothesis that the lexical and semantic properties of verbhood and nounhood are secondary and are primed and ultimately determined by their discourse roles, i.e. the determinants of nounhood and verbhood are predominantly pragmatic. The proposal the authors make is that linguistic entities set out as acategorical elements:

the continua which in principle begin with acategoriality, and which end with fully implemented nounhood or fully implemented verbhood, are already partly traversed for most forms. In other words, most forms begin with a propensity or predisposition to become Ns or Vs; and often this momentum can be reversed by only special morphology. It nonetheless remains true that this predisposition is only a latent one, which will not be manifested unless there is pressure from the discourse for this to occur (Hopper and Thompson 2004: 287).



In terms of its part-of-speech system Modern English, according to Vogel (2000: 263), has undergone a “degrammaticalisation shift from a specialised noun-verb language (with a grammaticalised part-of-speech system) towards a flexible type-token language (without a grammaticalised part-of-speech system).” For the more conservative, Vogel suggests that English might be thought of as having two parallel part-of-speech systems. “Thus, there are now two overlapping systems: a specialized noun-verb-adjective-adverb-system and a flexible noun/verb/adjective-adverb-system” (ibid. 277), which she illustrates in the following table:

Specialised	V	N	Adj	Adv
Flexible	V/N/Adj			Adv

Fig. 9. The Modern English part-of-speech system (Vogel 2000: 277).

We claim that it is the flexible system that is utilized in compounding.

And if the above arguments concern categorial assignment in general, Bauer’s (2001) idea of formal isolation as a basic criterion for compoundhood is open to interpretation and allows for acategorial treatment of compound-internal constituents.

Compound is a lexical unit made up of two or more elements, each of which *can* function as a lexeme independent of the other(s) in other contexts, and which shows *some* phonological and/or *grammatical isolation from normal syntactic usage* (Bauer 2001: 695, emphasis added).

Such an understanding of the constituents of a CV helps overcome the controversies involved in their classification, most of which arise from the dichotomization of the coordinate/subordinate distinction.

### 4.3 Standard classifications

Lieber<sup>11</sup> (2009: 360-361) in characterizing compounds in English, defines *stir-fry*, *trickle-irrigate*, *slam-dunk* and *blow-dry* as simultaneous endocentric coordinate compounds; *headhunt*, *machine-wash* and *spoon-feed* as endocentric verb-containing subordinate CVs, constituting “a marginal class”. She goes on to mention as endocentric compounds V+V patterns, which are “not freely formed” (ibid.: 359), *blow-dry*, *trickle-irrigate* and *slam-dunk*. The author expressly distinguishes between coordinate and subordinate CV types – i.e. *stir-fry* as opposed to *spoon-feed*. This classification is based on an implicit firm belief that the input elements into CVs are fully-fledged lexical items with specified part-of-speech membership. Only such an assumption can explain the forceful opposition between *blow-dry* and *stir-fry* as simultaneous coordinate CVs and *spoon-feed* and *machine-wash* as subordinate endocentric ones. The implicit assumption

<sup>11</sup> The classifying system of compounds that Lieber uses is the one presented by Bisetto and Scalise (2009) with a slight broadening of the subordinate class to include ones with subject-oriented interpretations.

that when we combine two linguistic entities with similar properties and level of granularity (verbal lexemes in our case), they are usually in a coordinate relationship, while combined linguistic entities with presumed differential properties are expected to contract syntactically motivated relations (in our case a subordinate relation between a verb and a noun interpreted as instrument within the frame of the verb).

Counter Lieber, Bauer believes that “[t]o trickle-irrigate is to irrigate in a particular way” (Bauer n.d: 8). In a same manner

[a]ccording to Carstairs-McCarthy (2002: 61) *stir-fry* and *freeze-dry* are both headed compounds. The hyponymy test works well here: stir-frying is a kind of frying (not a kind of stirring) and freeze-drying is a kind of drying (achieved by freezing). Again they seem to be excluded from the set of dvandvas (Bauer 2008: 4).

The subordinate vs. coordinate classification seems overdetermined, if we recognize the existence of well-elaborated word formation niches<sup>12</sup> based on sanctioned construction schemas.

By adopting ‘construction’<sup>13</sup> as an operationalizing tool for the idea of the continuity among all expressible resources in the symbolic inventory of language, we can postulate a generalized schema for *machine-wash*, *stir-fry* and *tumble-dry*, which will collapse the subordinate vs. coordinate divide determined to a large extent by the lexeme-status of input elements<sup>14</sup>.

[[P] Xi, j, k [R] Va, b, c ] V<sub>l</sub>, m, n → [to *process* something to a certain effect by exploiting SEMi, j, k effect]

X – compound internal constituents

P and R – sound sequences

j, a, m – lexical indexes specifying possible homonymous correlations with simplex lexemes.

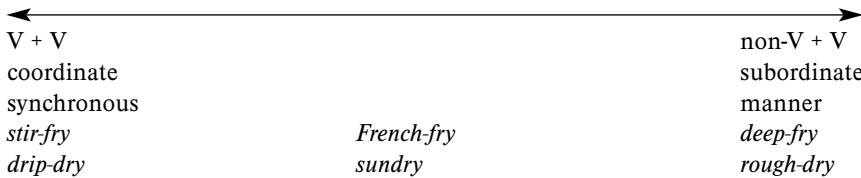
On the surface it seems that we have coordinate CVs where the lexical input structure suggests two verbs profiled in a relational concept. This interpretation sounds convincing for certain members of the word-formation niche – *stir-fry* might be interpreted as a synchronous coordinate V+V compound, but *deep-fry* does not yield such an analysis because there is no \*to deep verb in English and second because the more plausible semantic interpretation of the whole - *fry* niche tends

<sup>12</sup> Hüning (2009: 183) contends that “word formation processes often show semantic fragmentation: in the course of time they develop ‘semantic niches’, i.e. groups of words (subsets of a morphological category) kept together by formal and semantic criteria and extendable via analogy.”

<sup>13</sup> The construction schema (Booij 2005, 2007, 2009, 2010; Tuggy 2005 a,b; Langacker 1987; Lampert and Lampert 2010) is “a cognitive representation comprising a generalization over perceived similarities among instances of usage” resulting from “repeated activation of a set of co-occurring properties” (Barlow and Kemmer 2000: xxiii).

<sup>14</sup> The nature of the construction and its representation follows Booij (2010), with the distinction that in the current argumentation the X and V constituents are not of determinate lexical status.

towards the manner subordinate extreme of the cline. The generalized meaning is [TO FRY IN A CERTAIN WAY, I.E. TO A CERTAIN EFFECT] and it is only a secondary or contingent fact that some of the effects involve naming by a verb-like construction-internal constituent. It turns out that even generalizations involving a local niche-internal classification require the postulation of a cline which can pay justice to all the facts.



The cline captures the *-fry* niche with the generalized meaning [COOK TO A CERTAIN EFFECT]. *Stir-fry* will occupy the leftmost area of the cline, while *deep-fry* will indicate the rightmost extreme.

#### 4.4 Conceptualization-based classification

In the context of drawing parallels and contrasts between noun and verb classifications, McGregor (2002) postulates a difference between classes and categories based on the types of relations among the members thereof and argues for a distinction between superclassification and subclassification.

Certain noun-verb compounds in English (e.g. hand-pick, pistol-whip, horse-whip, test-drive, etc.) also represent a type of verbal subclassification: they specify subtypes of the event denoted by the verb. Gooniyandi, by contrast, shows a system of verb superclassification (McGregor 2002: 5).

A further discontinuity is defined between

- grammatical classification: systems of overt or covert classification of lexemes; and
- epistemological classification: systems of linguistic units that categorise a domain of (conceptual) referents0 (McGregor 2002: 22)

As can be surmised from the author's suggestions, certain CVs in English are instances of verbal subclassification, resembling the endocentric modifier nominal type (hyponymy-based). By implication it can be concluded that other CVs in English do not belong to the subclassifying type. Viewed from the perspective of epistemological classification, the other CVs belong to a superclassification class in which members create new individuated types of activities (conceptualizations), i.e. names of socio-culturally significant activities, usually pragmatically primed (e.g. *deadpan*, *fast-talk*, *hightail*).

CVs in English constitute a constructional idiom associated with two construal prototypes which enclose a gradient space of conceptualizations. The constructional

idiom can be represented as [X Y]v, where a subdivision is established between CVs in which Y can be associated via sub-classification with a homonymous simplex verb (A) e.g. *spray-paint, spoon-feed, headhunt*; and epistemologically novel ones where the second constituent might be associated with a simplex V or not (B) e.g. *hag-ride, to moonlight, to graymail, to sandbag, to blackball, to honeymoon*.

The first prototype actualizes the subclassification class of CVs and is represented by [X v]v where internal v is not categorically specified within the construction but is homonymous with a simplex verb.<sup>15</sup> The radial network of this prototype contains two active zones in which the following can be distinguished:

- a. [SPATIAL SPECIFIER v]v – Marchand’s genuine CVs: e.g. *outnumber, undergo, oversee*;
- b. [X F,D,P]v where F, D, P stand for lexically but not categorically specified inputs (e.g. *fry, dry, paint, etc.*) which is associated with well-elaborated word formation niches, e.g. *deep-fry; drip-dry; spray-paint*.

The second prototype constitutes the class of superclassification CVs and can be represented by [X Y]v, where the overall meaning of the CV involves the use of initial situational interpretation which provides the onomasiological motivation for the CV, e.g. *sandbag, piggyback, badmouth*.

## 5. CV SEMANTICS

Compositionality<sup>16</sup> in the lexical semantics of CVs arises out of the contribution of:

- i. the construction as a whole [X Z]v contributes the categorical specification and the instruction that the whole conceptual complex is to be profiled as a relational concept
- ii. the blending<sup>17</sup> of two input frames.

The specific lexical semantics of CVs involves the activation of two input conceptual frames which interact in specifiable ways. There are three different ways in which the frames can interact, interrelated with the ramifications of the specified radial networks of the two CV prototypes (as defined in 3):

---

<sup>15</sup> This restriction holds true throughout the inheritance hierarchy.

<sup>16</sup> Compositionality is to be understood as a non-trivial and non-standard compositional theory likely to capture the general way in which semantic parts of a compound configure into a semantic whole (Bundgaard et al. 2006, 2007). Compositionality constrains the construal of the overall compound signification.

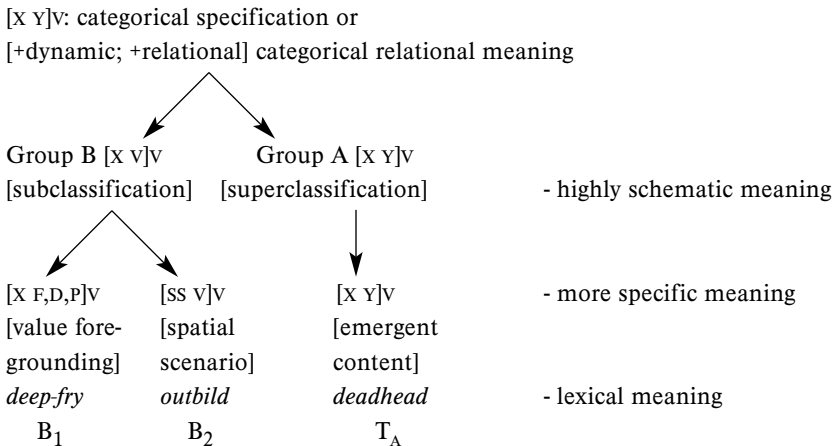
<sup>17</sup> Blending (involving both conceptual disintegration (Bache 2005, Hougaard 2005) and conceptual integration) is a powerful cognitive processing routine which functions as guiding context for constraining and establishing a semantic bond (Fauconnier and Turner 2002, Coulson 2001, Bache 2005, Bundgaard et al. 2006).

A. [X F,D,P]V – a single-scope blending of the two frames where the first frame fills an available slot in the second one and is foregrounded in a newly perspectivized profile of the second frame (e.g. *bottle-feed*, *spray-paint*, *cartwheel*)

B. [SPATIAL SPECIFIER V]V – blending where the two frames merge and the SPATIAL SPECIFIER frame augments the second frame by embedding it in a spatial scenario via the LOCATION branch of the EVENT metaphor. In such CVs the meaning is established on the basis of an asymmetrical double-scope network of frame blending (e.g. *overcome*, *under-rate*, *outnumber*)

[X Y]V – symmetrical double-scope blending where the generic space is a newly emergent one in which the features to be projected from the two input frames are selected in keeping with the graded salience hypothesis (Giora 1997, 2002; Huang 2009) and in which pragmatically driven mapping principles of relevance motivate the relevant elaboration of the blend.

These three frame-blending mechanisms correlate with the three readily identifiable zones in the conceptualization-based classificatory space of CVs:



Each CV realizes the merging of two input frames evoked by the compound-internal constituents and surfaces as a unified frame with

“a. A word sense’s **semantic frame** (what the word “means” or “evokes”) = **profile + background frame**

b. A word sense’s **profile**: what the word designates, asserts

c. A word sense’s **background frame**: what the word takes for granted, presupposes” (Goldberg 2010: 39)

The only constraint for the sanctioning of a CV schema of any of the lower levels the only constraint is for the conceptual content evoked by it to be construable as a single frame, which is fully congruent with Goldberg’s (2010: 39) argument that “[t]o

count as “coherent” and “individuable,” the situation or experience must be construable as a unit.”

The [X F,D,W]V, SINGLE SCOPE BLENDING PATTERN was illustrated in 4.3., so only the other two patterns are discussed below.

The [SPATIAL SPECIFIER V]V, ASYMMETRICAL DOUBLE-SCOPE BLENDING PATTERN leads to asymmetrical conceptual blending in which one of the outputs projected as a feature is blended as a default value in the meaning construction of the resultant CV. Being cognitively related to the homonymous preposition, these non-verbal constituents in CVs, just as the more abstract non-spatial meanings of prepositions, “tend to be derived from concrete, spatial senses by means of generalization or specialization of meaning or by metonymic or metaphoric transfer” (Cuyckens and Radden, 2002: xiii). In *outnumber* the input space of *out* projects the feature [BEYOND CERTAIN LIMITS], which is derived from [LEAVING A CONTAINER] into the composition in the generic space which combines with the meaning of [TO MAKE A TOTAL; REACH AN AMOUNT] as a default. Through completion, it also projects the constructional requirement that the blended space contain a counterpart to the agent that performs the verbal activity in the conceptual space in the verbal input so that the [LIMIT] meaning component can be set up by the emergent contrast between the counterparts in the running of the blend. The frame of *out* introduces a CONTAINER and imposes a spatial reading. VERTICALITY and PATH accompany this basic image-schema grounded metaphoric projection. VERTICALITY arises in the conflation of the two frames, the second of which involves QUANTITY IS VERTICAL MOVEMENT. Important for the configuring of the overall lexical meaning is also the FAR-NEAR schema. Besides introducing the CONTAINER schema, *out* implies MOVEMENT OUT OF THE CONTAINER, i.e. has a PATH/DIRECTIONAL component of meaning. Implicationally the FAR-NEAR schema will presuppose lack of control and will cancel the ‘exceed’, ‘be more powerful’ meaning of *outnumber*. However the VERTICALITY schema and the metaphor QUANTITY IS VERTICAL ELEVATION in the configured conceptualization override the far-near schema and associate power and control with vertical elevation. So the configuration involves QUANTITY IS VERTICAL ELEVATION conceived of as UPWARD MOVEMENT OF LEAVING A CONTAINER with the implication of control of the moving entity over the one remaining in the container<sup>18</sup>.

The verbs actualising the [X Y]V, SYMMETRICAL DOUBLE-SCOPE BLENDING PATTERN are invariably semantically exocentric as none of the constituents has any prevalence or leading role in profiling the blend. “Exocentricity<sup>19</sup> is an «anomaly» in language

---

<sup>18</sup> This analysis is based on the ideas of Lakoff and Johnson 1980, 1999; Evans and Green 2006, Evans and Tyler 2004, 2005; Tyler and Evans 2003, Cervel 2004.

<sup>19</sup> Benczes’s (2004, 2006) detailed analysis of creative N+N compounds in English reveals that creativity is synonymous with degree of extension and that they correlate with transparency of meaning (Benczes 2006: 189). Respectively semantic exocentricity corresponds to higher degree of creativity.

design in the following sense: describing a construction as exocentric means acknowledging that we cannot account for all the information conveyed by it” (Scalise and Guevara 2006: 185). The criterion of semantic exocentricity adopted here is understood as defined by Scalise et al. (2009: 59-60). “the third sense of exocentricity, in which the semantic class denoted by the compound cannot be predicted from the semantic class of their constituents.”

In such CVs both the constituents and the underlying constructional schema are co-activated as salient components of the blend. Consequently elaborate processes of metaphonymic (Goossens 2003) semantic creativity are involved in blending the two input frames so as to preserve the constructional requirement for a fleshed-out and richly emergent relational concept. *Hag-ride* is used as an illustrative example:

hag (input 1)

witches ride on broomsticks

witches torment people (including by riding them: sitting on their chest at night casting bad spells)

witches can bring bad luck

ride (input 2)

to move along in any way; be carried or supported

to sit on and manage a horse or other animal in motion; be carried on the back of an animal

Generic space

- RIDING IS TORMENTING

In the initial component of the blend the metonymy CAUSE FOR EFFECT is activated, while in the second component MEMBER OF A CATEGORY FOR THE CATEGORY operates. Riding is associated with LIFE IS A JOURNEY via MEMBER OF A CATEGORY FOR THE CATEGORY; then CONTROLLED FOR CONTROLLER is activated to reverse the cause-effect directionality of the riding schema; then the blend runs BEING RIDDEN BY A HAG IS LIVING IN A TORMENTED STATE.

## 6. FINAL COMMENTS

The peculiarities involved in configuring the meaning of a CV derive from the fact that “the conceptual base that underlies their predication is complex: ... a *complex scene*”<sup>20</sup> (Langacker 1987: 141; emphasis in the original). The cognitive-functional rationale of CVs in languaging<sup>21</sup> lies in their ability to compress situational scripts

---

<sup>20</sup> Langacker defines the semantic pole of an idiom in this quotation, but taking the liberty to use it in relation to compound verbs is justified by its appropriateness and aptness for describing the conceptual constitution of a compound verb.

<sup>21</sup> The term languaging is used as a shorter version of Slobin’s (2003) ‘thinking for speaking’, though the original more comprehensive term used by Maturana (1978) is also intended here.

(clause-like descriptions) into naming units whose constructional uniformity is guaranteed by overall verbal profiling, or actionalization construal (Langacker 2008). CVs rank high as convenient rich conceptual integration networks, as shorthand for events which enclose a cline between naming (thing/relation profiling) and description (situation profiling).

The tradeoff between explicitness and guided implicitness in the meaning constitution of CVs positions them as linguistic entities somewhere between explicit descriptions (situations) and implicitly construed names (relational profiling). This accounts for their cognitive power as construal mechanisms.

Just as the categorial, formal, semantic and origin properties of compounds represent a tightly interwoven totality, thus our analyses should try to keep as flexible and open an eye as possible without imposing dichotomous analytical categories which violate the natural fluidity of CVs as unique constructions in the hierarchical patterning of constructions in the fabrics of language.

## **BIBLIOGRAPHY**

### **Primary sources**

All examples of CVs have been taken from a corpus of CVs excerpted from the following:

### **Reference materials**

British National Corpus. 24 June 2011. <http://www.natcorp.ox.ac.uk/>.

Cambridge Advanced Learner's Dictionary Online. 27 June 2011. <http://dictionary.cambridge.org/>.

DAVIES, Mark. (2008-) The Corpus of Contemporary American English (COCA): 425 million words, 1990-present. 29 June 2011. <http://www.americancorpus.org>.

Dictionary.com Unabridged (Based on the Random House Dictionary, © Random House, Inc. 2011). 20 June 2011. <http://dictionary.reference.com/>.

FrameNet. 25 June 2011. <http://framenet.icsi.berkeley.edu/>.

*Longman Dictionary of Contemporary English* (2005). Pearson Education Limited.

*Oxford advanced learner's dictionary*. (2005). (7<sup>th</sup> ed.) Oxford: Oxford University Press.

*Oxford English Dictionary* Second Edition on CD-ROM (v. 4.0.0.3). (2009) Oxford University Press.

urbandictionary.com. 29 June 2011. <http://www.urbandictionary.com/>.

### **Academic papers and books**

BAUER, Laurie (1983) *English word-formation*. Cambridge: Cambridge University Press.

CLARK, Eve/Herbert CLARK (1979) "When Nouns Surface as Verbs." *Language* 55/4: 767-811.

ERDMANN, Peter (2000) "Compound verbs in English: Are they pseudo?" In: G. Tops/B. Devriendt/S. Geukens (eds.), 239-252.

HALL, Robert (1956) "How we noun-incorporate in English." *American Speech* 31/2: 83-88.



- MARCHAND, Hans (1969) *The Categories and Types of Present-Day English Word-Formation*. 2<sup>nd</sup> edition. München: Beck.
- MARTSA, Sándor (2007) "Construction of Meaning during Conversion." 7 May 2011. [husse-esse.hu/wp-content/.../construction-of-meaning-during-conversion.doc](http://husse-esse.hu/wp-content/.../construction-of-meaning-during-conversion.doc).
- MINKOVA, Donka/Robert STOCKWELL (eds) (2002) *Studies in the history of the English language: a millennial perspective*. Berlin: Mouton de Gruyter.
- PENNANEN, Esko (1966) *Contributions to the Study of Back-Formation in English*. Tampere: Acta Academiae Socialis. A. Vol.4.
- TOPS, Guy/Betty DEVRIENDT/Steven GEUKENS (eds) (2000) *Thinking English grammar to honour Xavier Dekeyser, Professor Emeritus*. Peeters.
- WALD, Benji/Lawrence BESSERMAN (2002) "The emergence of the verb-verb compound in twentieth century English and twentieth century linguistics." In: D. Minkova/P. Stockwell (eds), 417-447.

### **Fiction**

- CLARKE, Stephen (2005) *A Year in the Merde*. Black Swan.
- CLARKE, Stephen (2006) *Merde Actually*. Black Swan.
- CLARKE, Stephen (2008) *Merde Happens*. Black Swan.
- CLARKE, Stephen (2009) *Dial M for Merde*. Black Swan.
- CLARKE, Stephen (2011) *A Brief History of the Future*. Black Swan.
- CONROY, Pat (1996) *Beach Music*. Bantam
- COUPLAND, Douglas (2010) *Player One: What Is to Become of Us*. House of Anansi Press

### **References:**

- AARTS, Bas/David DENISON/Evelien KEIZER/Gergana POPOVA (eds) (2004) *Fuzzy Grammar: A Reader*, Oxford: Oxford University Press.
- ACKEMA, Peter/Ad NEELEMAN. 2004 *Beyond Morphology: Interface Conditions on Word Formation*. Oxford University Press.
- ANDERSON, Stephen (1992) *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- BACHE, Carl (2005) "Constraining conceptual integration theory: Levels of blending and dis-integration." *Journal of Pragmatics* 37, 1615–1635.
- BARLOW, Michael/Suzanne KEMMER (eds) (2000) *Usage Based Models of Language*. Stanford, CA: CSLI Publications.
- BARLOW, Michael/Suzanne KEMMER (2000) "Introduction: A usage based conception of language." In: BARLOW, M./S. KEMMER (eds), vii-xxviii.
- BAUER, Laurie (n.d.) "English Exocentric Compounds« (the English version of the paper published in French 'Les composés exocentriques de l'anglais')." In: D. Amiot (ed), *La composition dans uneperspective typologique*. Arras: Artois Presses Université, pp. 35-47). 20 September 2010. <http://www.victoria.ac.nz/lals/staff/Publications/Bauer-English-ExocentricCompounds.pdf>.
- BAUER, Laurie (2001) "Compounding." In: M. Haspelmath./E. König/W. Oesterreicher/W. Raible (eds), 695-707.
- BAUER, Laurie (2008) "Dvandva." *Word Structure*. Vol. 1: 1-20.

- BENCZES, Réka (2004) "Analysing exocentric compounds in English: a case for creativity." *The Even Yearbook 6 ELTE SEAS Working Papers in Linguistics*.
- BENCZES, Réka (2006) *Creative Compounding in English*. Amsterdam/Philadelphia: John Benjamin Publishing Company.
- BISETTO, Antonietta/Sergio SCALISE (2009) "The Classification of Compounds." In: R. Lieber/P. Štekauer (eds), 34–53.
- BOOIJ, Geert (2005) "Compounding and derivation: evidence for construction morphology." In: W. Dressler/D. Kastovsky/O. Pfeiffer/F. Rainer (eds), 109–132.
- BOOIJ, Geert (2007) *The Grammar of Words*. Oxford: Oxford University Press.
- BOOIJ, Geert (2009) "Lexical integrity as a formal universal: a constructionist view." In: S. Scalise./E. Magni/A. Bisetto (eds), 83–100.
- BOOIJ, Geert (2010) *Construction Morphology*. Oxford: Oxford University Press.
- BUNDGAARD, Peer/Svend OSTERGAARD/Frederik STJERNFELT (2006) "Waterproof Fire Stations? Conceptual schemata and cognitive operations involved in compound constructions." *Semiotica* 161/1-4, 363-393.
- BUNDGAARD, Peer/Svend OSTERGAARD/Frederik STJERNFELT (2007) "Meaning construction in the production and interpretation of compounds is schema-driven." In: *Acta Linguistica Hafniensia* 39:155-177.
- CERVEL, Sandra Peña. 2004. "The image-schematic basis of the event structure metaphor." In: *Annual Review of Cognitive Linguistics* 2: 127-158.
- CLARK, Eve/Herbert CLARK (1979) "When Nouns Surface as Verbs." *Language* 55/4: 767-811.
- COULSON, Seana (2001) *Semantic Leaps. Frame-Shifting and Conceptual Blending in Meaning Construction*. Cambridge: Cambridge University Press.
- CROFT, William/Allan CRUSE (2004) *Cognitive Linguistics*. Cambridge: Cambridge University Press.
- CRUSE, Alan (2004) *Meaning in language. An introduction to semantics and pragmatics*. Oxford: Oxford University Press.
- CUYCKENS, Hubert/Walter DE MULDER/Tanja MORTELMANS (eds) (2004) *Adpositions of Movement. Belgian Journal of Linguistics* 18. Amsterdam and Philadelphia: John Benjamins.
- CUYCKENS, Hubert/Günter RADDEN (2002) "Introduction." In: H. Cuyckens./G. Radden (eds), ix-xvii.
- CUYCKENS, Hubert/Günter RADDEN (eds) (2002) *Perspectives on prepositions*. Tübingen: Niemeyer
- DIRVEN, René/Ralf PÖRINGS, (eds) (2003) *Metaphor and Metonymy in Comparison and Contrast*. Berlin: Mouton de Gruyter.
- DIRVEN, René/Marjolijn VERSPOOR (2004) *Cognitive Exploration of Language and Linguistics*. Amsterdam/Philadelphia: John Benjamins.
- DRESSLER, Wolfgang/Dieter KASTOVSKY/Oskar PFEIFFER/Franz RAINER (eds) (2005) *Morphology and its demarcations*. Amsterdam/Philadelphia: John Benjamins.
- EVANS, Vyvyan/Andrea TYLER (2004) »Rethinking English Prepositions of Movement: "The Case of *To* and *Through*." In: H. Guyckens./W. DE Mulder/T. Mortelmans (eds), 247-270.

- EVANS, Vyvyan/Andrea TYLER (2005) "Applying cognitive linguistics to pedagogical grammar: The English prepositions of verticality." *Revista Brasileira de Linguística Aplicada*, 5 (2): 11-42. 18 June 2010. <http://www.vyvevans.net/Applycoglxpedagogy.pdf>.
- EVANS, Vyvyan/Melanie GREEN (2006) *Cognitive Linguistics: An Introduction*, Edinburgh University Press.
- EVANS, Vyvyan (2006) "Lexical Concepts, Cognitive Models and Meaning-Construction." *Cognitive Linguistics* 17/4, 491-534.
- EVANS, Vyvyan (2007) "Towards a cognitive compositional semantics: An overview of LCCM Theory." In: U. Magnusson./H. Kardela/A. Glaz (eds), 11-42.
- EVANS, Vyvyan (2009) *How words mean: Lexical Concepts, Cognitive Models and Meaning Construction*. Oxford: Oxford University Press.
- EVANS, Vyvyan/Stéphanie POURCEL (eds) (2009) *New Directions in Cognitive Linguistics*, Amsterdam/Philadelphia, 'Human Cognitive Processing' 24, John Benjamins Publishing Company.
- FARRELL, Patrick (2001) "Functional Shift as Category Underspecification." *English Language and Linguistics* 5.1, 109-130.
- FAUCONNIER, Gilles/Mark TURNER (1998) "Conceptual integration networks." *Cognitive Science* 22/2. 133-187.
- FAUCONNIER, Gilles/Mark TURNER (2002) *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.
- FILLMORE, Charles (2006) "Frame Semantics." In: GEERAERTS, D. (ed), 373-400.
- GEERAERTS, Dirk/Hubert CUYCKENS (eds) (2007) *The Oxford Handbook of Cognitive Linguistics*. Oxford: Oxford University Press.
- GEERAERTS, Dirk (ed) (2006) *Cognitive Linguistics: Basic Readings*. Berlin/New York: Mouton de Gruyter.
- GENTNER, Dedre/Susan GOLDIN-MEADOW (eds) (2003) *Language in Mind: Advances in the Study of Language and Thought*. Cambridge, Massachusetts: The MIT Press.
- GIORA, Rachel (1997) "Understanding figurative and literal language: the Graded Salience Hypothesis." *Cognitive Linguistics* 7/1, 183-206.
- GIORA, Rachel (2002) "Literal vs. Figurative Language: Different or Equal?" *Journal of Pragmatics* 34: 487-506.
- GOLDBERG, Adele (2010) "Verbs, constructions and semantic frames." In: M. Rappaport Hovav./E. Doron/I. Sichel, (eds), 39-58.
- GOOSSENS, Louis (2003) "Metaphtonymy: The interaction of metaphor and metonymy in expressions for linguistic action." In R. Dirven/R. Pörings (eds), 349-377.
- GUEVARA, Emiliano/Sergio SCALISE (2004) "V-compounding in Dutch and Italian." 27 June 2011. <http://morbo.lingue.unibo.it/user-documents/guevara/nl-it-24-02-04-final-format.pdf>
- GUEVARA, Emiliano/Sergio SCALISE (2009) "Searching for Universals in Compounding." In: Scalise, S./E. Magni/A. Bisetto (eds), 101-129.
- HASPELMATH, Martin/Ekkehard KÖNIG/Wulf OESTERREICHER/Wolfgang RAIBLE (eds) (2005) *Language Typology and Language Universals*. Berlin/New York: Walter de Gruyter.
- HOPPER, Paul/Sandra THOMPSON (2004) "The discourse basis for lexical categories in Universal Grammar." In: B. Aarts./D. Denison/E. Keizer/G. Popova (eds), 247-291.

- HOUGAARD, Anders (2005) "Conceptual disintegration and blending in interactional sequences: A discussion of new phenomena, processes vs. products, and methodology." *Journal of Pragmatics* 37 (10), 1653 – 1685.
- HUANG, Mimi (2009) "Solving the riddle of metaphor: A salience-based model for metaphorical interpretation in a discourse context." In: V. Evans./S. Pourcel (eds), 107-126.
- HÜNING, Matthias (2009) "Semantic niches and analogy in word formation: Evidence from contrastive linguistics." In: *Languages in contrast*. Vol. 9, No. 2. 183-201.
- JACKENDOFF, Ray (2002) *Foundations of Language*. Oxford: Oxford University Press.
- JACKENDOFF, Ray (2009) "Compounding in the Parallel Architecture and Conceptual Semantics." In: R. Lieber./P. Štekauer (eds), 105-128.
- JACKENDOFF, Ray (2010) *Meaning and the Lexicon: The Parallel Architecture 1975–2010*. Oxford: Oxford University Press.
- JESPERSEN, Otto (1954) *A Modern English Grammar on Historical Principles*. London: Bradford and Dickens.
- KIEFER, Ferenc (1993) "Thematic Roles and Compounds." *Folia Linguistica* XXVII/1-2, 45-55.
- LAKOFF, George/Mark JOHNSON (1980) *Metaphors We Live By*. Chicago: Chicago University Press.
- LAKOFF, George/Mark JOHNSON (1999) *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. Basic Books.
- LAMPERT, Martina/Günther LAMPERT (2010) "Word-formation or word formation? The formation of complex words in Cognitive Linguistics." In: A. Onysko./S. Michel (eds), 29-73.
- LAMPERT, Martina (2009) *Attention and Recombinance. A Cognitive-Semantic Investigation into Morphological Compositionality in English*. Peter Lang. Frankfurt am Main.
- LANGACKER, Ronald (1987) *Foundations of Cognitive Grammar*, vol. 1. Stanford: Stanford University Press.
- LANGACKER, Ronald (1990) *Concept, Image, Symbol*. Berlin: Mouton.
- LANGACKER, Ronald (2008) *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.
- LIEBER, Rochelle (1981) "Morphological Conversion within a Restrictive Theory of the Lexicon." In: M. Moortgat/H. Hulst/T. Hoekstra (eds), 161-200.
- LIEBER, Rochelle (2009) "IE, Germanic: English." In: R. Lieber/P. Štekauer (eds), 357 - 370.
- LIEBER, Rochelle/Pavol ŠTEKAUER (eds) (2009) *The Oxford Handbook of Compounding*. Oxford University Press
- MAGNUSSON, Ulf/Henryk KARDELA/Adam GLAZ (eds) (2007) *Further Insights into Semantics and Lexicography*. Lublin, Poland: Wydawnictwo UMCS.
- MARCHETTI, Giorgio (2006) "A Presentation of Attentional Semantics." *Cognitive Processing* 7, 163–194.
- MASINI, Franchesca (2009) "Phrasal lexemes, compounds and phrases: A constructionist perspective." *Word structure* 2/2, 254–272.
- MATURANA, Humberto (1978) "Biology of Language: The Epsitemology of Reality." In: G. Miller./E. Lenneberg (eds), 27-63.
- MCGREGOR, William (2002) *Verb Classification in Australian Languages*. Mouton de Gruyter. New York.

- MICHAELIS, Laura/Knud LAMBRECHT 1996. "Toward a construction-based theory of language function: The case of nominal extraposition." *Language* 72, 215-247.
- MILLER, George/Elizabeth LENNEBERG (eds) (1978) *Psychology and Biology of Language and Thought: Essays in Honor of Eric Lenneberg*. New York: Academic Press.
- MOORTGAT, Michael/Harry HULST/Teunen HOEKSTRA (eds) (1981) *The Scope of Lexical Rules*. (Linguistic Models 1). Dordrecht: Foris Publications.
- NAGANO, Akiko (2007) "Marchand's analysis of back-formation revisited: back-formation as a type of conversion." *Acta Linguistica Hungarica*, 54/1, 33-72.
- NERLICH, Brigitte/Zazie TODD/Herman VIMALA/David CLARKE (eds) (2003) *Polysemy: Flexible Patterns of Meaning in Mind and Language*. Berlin: Mouton de Gruyter.
- ONYSKO, Alexander/Sascha MICHEL (eds) (2010) *Cognitive Perspectives on Word formation*. Berlin/New York: De Gruyter Mouton.
- RADDEN, Günter/Rene DIRVEN (2007) *Cognitive English Grammar*. John Benjamin Publishing House.
- RAPPAPORT HOVAV, Malka/Edit DORON/Ivy SICHEL, (eds) (2010) *Syntax, Lexical Semantics, and Event Structure*. Oxford: Oxford University Press.
- SCALISE, Sergio/Antonio FÁBREGAS/Francesca FORZA (n.d.) "Exocentricity in Compounding." 17 August 2009. <http://morbo.lingue.unibo.it/userdocuments/scalise/exocentricityincompounding.pdf>.
- SCALISE, Sergio/Elisabetta MAGNI/Antionietta BISETTO (eds) (2009) *Universals of language today*. Berlin: Springer.
- SCALISE, Sergio/Emiliano Guevara (2006) "Exocentric Compounding in a Typological Framework." *Lingue e Linguaggio* V.II, 185-206.
- SCALISE, Sergio/Elisabetta MAGNI/Antionietta BISETTO (eds) (2009) *Universals of language today*. Berlin: Springer.
- SCHMID, Hans-Jorg (2007) "Entrenchment, Salience, and Basic Levels." In: D. Geeraerts/H. Cuyckens (eds), 117-138.
- SCHMID, Hans-Jorg (2011 [2005]) *English morphology and word-formation: An introduction*. Transl. Ruth Owen and Sue Bollinger. Berlin: Erich Schmidt Verlag.
- SLOBIN, Dan (1987) "Thinking for speaking." *Proceedings of the Berkeley Linguistics Society* 13, 435-444.
- SLOBIN, Dan (2003) "Language and Thought Online: Cognitive Consequences of Linguistic Relativity." In: D. Gentner/S. Goldin-Meadow (eds), 157-192.
- TALMY, Leonard (2000) *Toward a cognitive semantics*. 2 volumes. Cambridge, MA: Massachusetts Institute of Technology.
- TALMY, Leonard (2007) "Attention Phenomena." In: D. Geeraerts/H. Cuyckens (eds), 264-293.
- TYLER, Andrea/Vyvyan EVANS (2003) "Reconsidering Prepositional Polysemy Networks: The Case of Over." In: B. Nerlich/Z. Todd/H. Vimala/D. Clarke (eds), 95-160.
- TUGGY, David (2005a) "Abrelatas and Scarecrow: Exocentric Verb-Noun Compounds as Illustrations of Basic Principles of Cognitive Grammar." *International Journal of English Studies*. 3/2 *Monograph: Contrastive Cognitive Linguistics*, 25-61.
- TUGGY, David (2005b) "Cognitive Approach to Word-Formation." In: R. Lieber/P. Štekauer (eds), 233-267.

- VAN DER ZEE, Emile/Urpo NIKANNE (2000) Introducing Cognitive Interfaces and Constraints on Linking Cognitive Information. In: E. Van Der Zee/U. Nikanne (eds), 1-17.
- VAN DER ZEE, Emile/Urpo NIKANNE (eds) (2000) *Cognitive Interfaces: Constraints on Linking Cognitive Information*. Oxford: Oxford University Press.
- VERHAGEN, Arie (2007) "Construal and Perspectivization." In: D. Geeraerts/ H. Cuyckens (eds), 48-82.
- VOGEL, Petra (2000) "Grammaticalization and Part of Speech Systems." In: P. Vogel/B. Comrie (eds), 259-284.
- VOGEL, Petra/Bernard COMRIE (eds) (2000) *Approaches to the Typology of Word Classes*. New York: Mouton de Gruyter.

#### Abstract

### BEYOND DICHOTOMIES: ON THE NATURE AND CLASSIFICATION OF COMPOUND VERBS IN ENGLISH

The study of compound verbs in English poses numerous problems, among which even their recognition as compounds on grounds of their derivation. Resulting from at least three different word-formation patterns, compound verbs constitute a heterogeneous class of complex lexemes. Their status as actual compound lexemes invites the differentiation between compounding as a word-formation process and compounds as a special class of lexemes. Even within the latter, compound verbs display marked properties at least in relation to the inability of standard classifications of compounds to capture and compromise their lexical uniformity and their heterogeneous origin. The adoption of a position in which it is argued that compound verbs in English constitute a constructional idiom and the application of scalar analytical notions which combine word-formationist and lexical-semantic accounts cast in the general framework of the cognitive linguistic enterprise yield informative generalizations concerning the linguistic and conceptual properties of compound verbs in English. In view of Radden and Dirven's (2007: 41-46) claim that we do not need "more than two basic types of conceptual units things and relations" in order to establish linguistically relevant conceptual distinctions, compound verbs pose a problem for neat dichotomous treatment as they very often both conceptually and in terms of form include a "thing" (e.g. to *flat-hunt*, to *house-sit*, to *fellow-feel*, to *case-harden*, etc.) and thus come closer to a "situation" than to a "relation". Exactly because of the fact that compound verbs profile/perspectivize "situations" as "relations", they function as special construal mechanisms and as such do not fit the subordinate/coordinate distinction, because they *name* situations. In view of the above the paper treats compound verbs as a constructional idiom whose analysis necessitates the recognition of the role of conceptual conversion mechanisms, scalar classificatory and interpretative criteria and uniform lexico-semantic treatment.

Povzetek  
ONSTRAN DIHOTOMIJ: O NARAVI IN TIPOLOGIJI  
GLAGOLSKIH ZLOŽENK V ANGLEŠČINI

Pri proučevanju glagolskih zloženek v angleščini trčimo na številne probleme, med drugim tudi, kako določiti status zloženke zgolj na podlagi besedotvornega vzorca. Glagolske zloženke, ki lahko nastanejo vsaj iz treh različnih besedotvornih vzorcev, tvorijo heterogeno skupino večleksemskih enot. Prav specifičnost teh enot, da so dejansko zloženi leksemi, nas sili, da razlikujemo med zlaganjem kot besedotvornim postopkom in zloženkami kot specifično vrsto leksemov. Za glagolske zloženke kot posebno vrsto leksemov lahko ugotovimo, da imajo specifične lastnosti, tako da standardne tipologije zloženek ne morejo zajeti in zaobjeti njihove leksikalne enovitosti in raznolikosti načinov nastanka. Če sprejmemo stališče, da predstavljajo glagolske zloženke v angleščini konstrukcijske idiomatske enote, in če se odločimo za uporabo skalarnih analitičnih pojmov, ki združujejo besedotvorni in leksikalni pristop, in če vse naštetost vključimo v splošni okvir kognitivnega jezikoslovja, potem pridemo do zanimivih splošnejših veljavnih ugotovitev, ki zadevajo jezikovne in konceptualne posebnosti glagolskih zloženek v angleščini. Upoštevajoč trditev, ki jo lahko preberemo v Radden in Dirven (2007: 41-46), da za določanje jezikovno relevantnih konceptualnih razlik zadostujeta že dva osnovna »tipa konceptualnih enot, to so stvari in razmerja,« potem moramo ugotoviti, da predstavljajo glagolske zloženke poseben problem v prizadevanju po jasnem dihotomičnem razlikovanju, saj pogosto tako na konceptualni ravni kot formalno vključujejo »stvar« (npr. to *flat-hunt*, to *house-sit*, to *fellow-feel*, to *case-harden* ipd.) in so torej bližje »stanju« kot »razmerju«. Prav zato ker glagolske zloženke »stanje« profilirajo/zarisujejo kot »razmerje«, delujejo kot posebni konceptualni mehanizmi, ki jih kot takih ne moremo opisati z običajnim razlikovanjem med podredno in priredno strukturo, saj stanja *poimenujejo*. Izhajajoč iz povedanega, pričujoči članek glagolske zloženke obravnava kot konstrukcijske idiomatske enote, ki za to, da jih lahko ustrezno analiziramo, terjajo, da priznamo pomen konceptualnih konverzijskih mehanizmov, skalarnih tipoloških in interpretacijskih kriterijev in enovitega leksikalno-semantičnega pristopa.





## ISLAND MORPHOLOGY: MORPHOLOGY'S INTERACTIONS IN THE STUDY OF STEM PATTERNS

### 1. INTRODUCTION

Discussing *island* morphology, or morphology in itself, for that matter, entails the presupposition that morphology exists as a component in the grammar that has its own complexity and idiosyncrasies, and need not always interface and interact with others. The use of the metaphorical term *island morphology* should be equated to an autonomous level of morphology where complexity can be observed.<sup>1</sup>

The morphological complexity at issue here is the *redundant* type; that which makes no contribution to the form-meaning relation, and can be treated as a system of 'pure morphological functions' that have an independent 'life of [their] own' (Brown 2008: 6). The study of such functions, which has to do with the observation of how forms change, and what patterns result from this, which essentially have no value/relevance to the syntax and/or meaning, is referred to as *morphological features* (Corbett/Baerman 2006: 232). Asserting that such features have no interaction with syntax, the study will also probe the extent to which such features are allowed to border with phonology. As will be shown, the significance and autonomy of such features is best displayed in instances where the phonology seems to predict a pattern with which the morphology does not abide, thus resulting in a mismatch between the phonological and morphological forces.

What follows in section (2) is a discussion on morphological features, and their significance in a canonical typology account within the context of canonical paradigms. Section (3) briefly discusses the traditional distinction across verbs in the Semitic part of Maltese (henceforth SM), which will be the source of the paradigmatic data.<sup>2</sup> This is then followed by the verbal paradigmatic data which will show how syncretic stem indexations render patterns that may in turn vary across lexemes, thus creating

---

\* *Author's address:* University of Surrey, Faculty of Arts and Human Sciences, Surrey Morphology Group, Surrey GU2 7XH, United Kingdom. Email: maris.camilleri@surrey.ac.uk

<sup>1</sup> The term used, apart from metaphorically imparting the theoretical morphological claim in this paper, i.e. that morphology is autonomous, also extends itself nicely as a pun, considering that the data will be taken from Maltese; a language spoken in an island in the middle of the Mediterranean Sea.

<sup>2</sup> It is very much common to focus on either the Semitic part or the Romance part of the language when discussing morphology in Maltese. This is usually done in order to be able to capture any potential variation across the different elements in the language. After establishing how SM verbal morphology works in what will be looked at here, it would then be reasonable to do the same task for the Romance part and eventually compare the results.

*stem-based inflectional classes*; the subject matter at hand. Section (4) summarises the findings and concludes the paper.

## 2. MORPHOLOGICAL FEATURES

### 2.1 Form Classes

Morphological features are best illustrated through the study of *form classes*. The analysis of form-classes looks at output realisations; focusing solely on their surface structure (refer to Anderson's (2011) reply to Corbett (2011)), and observes the patterns formed across distinct segments, such as affixation, tone variation, and/or stem-forms. The study related with the autonomous morphological system displayed by the affixal material, or the relation of this with tone patterns, is referred to as the analysis of *inflectional classes*, while that involving stem-forms, or the relation of stem-forms with affixes and/or tone patterns, which create some sort of pattern in the paradigm's cells grouping, is referred to as *syncretic indexing* (Corbett 2004, 2007, 2009), (Baerman 2011), (Corbett/Baerman 2006, 2010), (Baerman/Corbett forthcoming). These features can be understood as distinct parameters of complexity. The former can be thought of as involving a horizontal parameter that looks across lexeme sets, and observes any comparisons/contrasts in the patterns displayed by the affixal material, while the latter, a vertical parameter that looks at the complexity which arises from within the paradigm. In both instances the patterns come about as a result of the different organisation of the array of inflected and/or lexical material in a given paradigm or across paradigms. Such organisatory patterns come about through different groupings of the surface forms as a result of their morphological sameness - *syncretism*, when these should in theory be distinct, since they are essentially the realization of distinct morphosyntactic features (Corbett 2005) (Baerman et al. 2005). The reason for assuming that such *form classes* are morphological features and hence irrelevant to the syntax (Alexiadou 2004: 21), is that, for example, the syntax does not care whether the particular affix which is being used is solely related with a set of lexemes and not another, as long as the correct word-form and realisation affix is selected to satisfy the morphosyntactic requirement.

The discussion here will pivot both the vertical and horizontal parameters mentioned above. On the vertical parameter we will observe the grouping of paradigmatic cells on the basis of the pattern rendered by the stem-form alternations in the different cells based on the syncretic stem index features. To this, a horizontal dimension will be added, which will illustrate the different paradigmatic stem-alternation patterns that cut across different lexical classes, which, as will be shown in section (3.2), may or may not be phonologically-conditioned, and whose distribution across the classes is also 'syntactically-unmotivated' (Corbett 2009: 5). The interaction of both the vertical and horizontal parameters will result in the analysis of *stem-based inflectional classes* (hereafter SBICs), which are taken to be analogically parallel to the *inflectional classes* mentioned above, but which are

derived from observing stem patterns instead of affixal material patterns<sup>3</sup>. In this study, stem-form patterns from SM will be looked at in relation to their interaction with affixes, as well as traditional verbal classifications, described in section (3.1), which rely heavily on an account based on a reference to the *underlying representation* (UR), consisting of a reference to an underlying triconsonantal root (C-root) morpheme. By virtue of this SBIC analysis of verbal paradigms in SM, verbs will be categorised systematically, not anymore on the basis of an abstracted UR, which may at times, not even be representative of the differences or similarities across verbal sets, but rather on the basis of their surface structure stem patterns and their distinct ways of interaction with other elements in the word-form, such as the inflectional affixes.

## 2.2 Stem patterns as a consequence of non-canonical paradigmatic effects

What is interesting about *form-classes* is that essentially these come about as a result of a non-canonical behaviour, or the interaction of non-canonical phenomena. The canonical approach to morphology can be understood as that manifestation of a morphological phenomenon/element, that unarguably represents that which best fits the set parameters or criteria (Corbett 2007: 9), which however need not be that which is prototypical, as it may even be non-existent (Corbett 2010: 3). Looking at that which is non-canonical means that one moves away from the canonical, and the further away one is from the canon, the more of a non-canonical instance it becomes (Corbett 2011: 2).

The non-canonical paradigmatic aspect which this study will be concerned with is internal non-canonicity, whereby while the lexical material in the paradigm i.e. the stem is expected to be the same; non-alternating (Corbett 2007: 10), this is not the case in non-canonical paradigms, giving rise to the presence of distinct stem-forms which may pattern differently across lexemes, thus creating SBICs by virtue of an independent organisatory system. The independent organisatory system is determining factor, since different SBICs may have the same number of stem-slots, but differ solely on the organisatory pattern. The data will also provide instances where there is an interaction of alternating stem-forms with other non-canonical phenomena such as syncretism. As will be shown, this will be a rather important interaction as it will account for the mismatches that may exist between the number of designated stem-slots, which comes about as a result of the syncretic stem indexation, vs. the actual number of stem-forms in the paradigm, and will be responsible for the internal SBIC stem pattern variation. The discussion in section (3.2.4) will also include a case illustrating the interaction of internal non-canonicity with suppletion, where the stem-form is not only altered, but involves a radical change.

Apart from being non-canonical because the stem alternates and creates distinct patterns, there is an additional non-canonical behaviour in SM verbal paradigms.

---

<sup>3</sup> Note that this implies that the view taken here is such that both affixes and stem-alternation patterns both constitute inflection in Maltese.

The stem may not simply bear lexical meaning alone, but is at times also endowed with grammatical features. This happens when in a SBIC, such as that discussed in section (3.2.4), the stem alternates on the basis of a coherent 1<sup>^</sup>2 vs. 3 PERSON feature distinction in the PERFect sub-paradigm, where the stem-forms used for the lexeme *zar* ‘visit’, for example, are *zor* and *zār* respectively. These features in turn overlap with those realised by the inflectional affixes (Baerman/Corbett 2011). This is regarded as additionally non-canonical because in ‘the canonical situation, lexical meaning (and only that) is conveyed by lexical material, the stem; grammatical meaning, and only that is conveyed by the inflection’ (Corbett 2011: 112).

### 2.2.1 Degrees of non-canonicity

Having established that the presence and formation of stem alternation patterns is non-canonical, it will be shown that there is a gradience of non-canonicity, such that for example, the more stem-form overlap across the PERFect-IMPERFect sub-paradigms, i.e. the more the syncretic interactions, which will result in an increased feature conflation on the stem, the more non-canonical the pattern is. The same follows for the number of members and/or stem patterns a SBIC takes. While a SBIC may be the most populous, thus potentially more canonical, among the other classes, it may still, on the other hand, be fragmented in distinct stem patterns, thus perhaps making it less canonical, in that it does not display uniformity across its members. Another parameter on which to grade the classes’ canonicity may potentially be related with the actual (in)stability of the SBIC’s members (refer to sections (3.2.2) and (3.2.4)), where it seems to be more canonical to have members participating in one SBIC only, rather than in two. All these parametric ideas can thus be further imposed on this non-canonical phenomenon to be able to capture the variation and the behaviour pertaining to these distinct SBICs.

## 3. THE DATA

### 3.1 Verbs in Semitic Maltese

The traditional classification of SM and Arabic verbs is that, on the basis of their UR, there is a: *Strong, deaf*, and *weak* verbal distinction; e.g. *kiber* ‘grow’, *mess* ‘touch’, and *mar* ‘go’ respectively. The *weak* class of verbs is further divided into three, depending on the position of the weak consonant (i.e. a *w* or *j*) in the UR; initial (*assimilative*) *wiret* ‘inherit’ (**w-r-t**), medial (*hollow*) *mar* ‘go’ (**m-w-r**), or final (*lacking/defective*) *qara* ‘read’ (**q-r-j**). Ryding (2005: 55) treats these distinctions as *inflectional classes*, however, as the reader must realise, this term has nothing to do with that used in section (2.1) above. *Inflectional classes*, as a term used in the wider literature, refers to the variation found in the realisation of affixal material across different classes of lexemes. This is not what we have here, however, as the inflectional material does not vary across these different types of verbs. What we do see however, in relation to the above verbal classifications is that at times, this distinction does in fact correlate with a distinct SBIC. Thus, for want of clarity, this term, referring to the traditional classification will

be abandoned, and the term used in Aquilina (1973: 144); *verbal bases*, will be used instead. This term is in Maltese traditionally used to refer collectively to the distinct types of verbs that exist in SM with reference to their URs, such as C-C-C (*strong*), w-C-C (*weak-initial*), C-C<sub>1</sub>C<sub>1</sub> (*geminate*) etc. The use of the term here, whilst retaining the same verbal classification understanding, will do this not on the basis of an UR, but rather on the stem-shape, such as CVCVC (which would include both *strong* and *weak-initial* verbs), CV:C (for *hollows*), and so on. In this paper it will be shown that the taxonomy of SM verbal bases on the basis of their UR does not always cross-cut/correlate with the classification that will be provided here, based on the observed surface stem-form patterns of the 1<sup>st</sup> *binyan* verbs. On the other hand, in some instances, membership in a SBIC does correlate exclusively with a particular verbal-base. We will also see that while traditional Maltese grammars always distinguish a handful of verbs as *irregular*, such as *ha* ‘take’ on the basis of an obsolete UR that includes the Arabic *alif* as the C-root’s first radical, it will be shown that in some cases, there is nothing that indicates, at least from the surface paradigmatic pattern that these should be distinguished from other *regular* verbs. A real instance of *irregularity* is, under the following classification, equated with a case of non-canonicity which follows from the discussion in section (2.2.1), where an irregular verb is understood as that which is a member of its own distinct SBIC or pattern. Thus, the classification of verbs, and the *regular/irregular* dichotomy will be motivated solely from that which can be observed from the surface paradigmatic patterns, based on a non-autonomous morphological system that seems to coordinate across surface forms through the creation of stem patterns, within and across paradigms.

### 3.2 The inventory of stem-based inflectional classes in SM

Before describing the SBICs, a brief overview of the data that will be used to illustrate the phenomenon will be given. The syncretic stem patterns which shall be looked at are derived from observations from the PERF-IMPERF sub-paradigm systems in SM. The PERF sub-paradigm in Maltese is characterised by suffixes which realise the conflated morphosyntactic features for PERS, NUMber, and ASPect, as well as GENDER where appropriate. In the IMPERF sub-paradigm prefixes realise PERS, ASP, and GEND, while suffixes realise NUM feature values. While the paradigm of a Maltese verb also includes an IMPERATIVE sub-paradigm, yet, this will not feature in the discussion here. Stem-alternations in SM verbal paradigms render their own stem-classes; i.e. different means with which stem-form alternations come about, which include syllable-structure type alternations, ablaut-changes, stem-extensions, and suppletion. While it is only the end result that we will be concerned with here, i.e. the stem-alternation in itself, whenever required there will be a reference to the actual mechanism which is responsible for this change.

Recall from section (2.2) that the criterion responsible for distinguishing among SBICs is the varied syncretic stem indexing pattern they employ. Thus, while the number of stem-slots can be the same across two classes, they have to differ by virtue of the stem indexing pattern involved, as is the case for SBIC I and III in sections (3.2.1) and

(3.2.3) respectively, which both have five stem-slots but whose syncretic pattern is grouped differently. We will also see how members of a given class need not be homogeneous, and may have varied stem patterns resulting from non-distinct stem-forms filling in separate stem-slots. Sections (3.2.1) till (3.2.6) below will illustrate and further discuss the individual six distinct SBICs in SM, and the variation involved. From the data which will be given below one would want to say that some classes come about as a result of a stem-alternation that is determined by morphotactic constraints that depend on the suffix's phonology, particularly in the PERF sub-paradigm. However, having said this, while the suffixes are generally invariant, we still get SBIC and stem pattern variation. Furthermore, if morphotactics determined patterns, then we would expect the stem-form used with V-initial suffixes in the IMPERF sub-paradigm to be always syncretic with the PERF stem-forms that are followed by a V-initial suffix themselves. This is not the case however. When mismatches of the sort arise, then we have a case of real morphological patterns which cannot be explained through any phonological considerations, thus highlighting the independent island morphology.

### 3.2.1 SBIC I

Stem-based inflectional class I is distinguished by having five stem-slots which require an alternating IMPERF sub-paradigm where the stem-pattern splits and is indexed for the NUM feature values SG vs PL; *lqot-* and *lqt-*, respectively in the *laqat* 'hit' paradigm, for example, and a conflated, non-motivated, non-coherent, morphomic split in the PERF, involving three distinct stem-slots that are indexed for 3SGM, 3SGF^3PL, and 1^2 (*laqat*, *laqt*, and *lqat*).<sup>4</sup> This class happens to be the most populous, including almost all the distinct types of verbal bases in SM apart from the *deaf* CVC<sub>1</sub>C<sub>1</sub> base (as displayed in sections (3.2.1.1-3.2.1.3)). The biggest set of paradigms however, belong to the traditional *strong* class with verbal-base C<sub>[-w/j]</sub>VCVC.<sup>5</sup> This SBIC will illustrate the stem pattern variation that can exist within the same class, where while being unified by the same number of stem-slots and the syncretic stem-pattern in the members' paradigms, varied patterns are brought about as a result of overlapping forms across the PERF-IMPERF paradigms. What is interesting to observe is that while the *strong* triliteral class involves six ablaut classes (section 3.2.1.1), the pattern of variation does not necessarily cross-classify with the distinct ablaut-classes, and there may be distinct sub-pattern variations within the same ablaut-class.<sup>6</sup>

---

<sup>4</sup> Note that, while one may wish to argue that this stem-slot distinction should be reinterpreted as 3SGM, 3, 1^2, with the latter two reflecting a distinction across speech act participants, yet, this 3PERS underspecification cannot hold, as there is nothing which requires that these are to be always conflated in that way, as shown in SBIC II, III, and VI.

<sup>5</sup> gh- including verbs will not be discussed in this study, as they involve a number of phonetic idiosyncrasies that need not concern us here, since they add nothing to our discussion here.

<sup>6</sup> An ablaut-class is here understood as the distinct vocalism pattern which refers to the two stem vowels (V1 and V2) of the PERF's 3SGM stem-form, which in Semitic languages is taken to be the citation form. Thus, the a-a ablaut class includes verbs such as *dahal* 'enter', and the e-e would include *qered* 'destroy'.

### 3.2.1.1 Strong verbal base ablaut classes

a-a		
	laqat 'hit'	
	PERF	IMPERF
1.SG	lqat-t	n-o-lqot
2.SG	lqat-t	t-o-lqot
3.SG.M	laqat	j-o-lqot
3.SG.F	laqt-et	t-o-lqot
1.PL	lqat-na	n-o-lqt-u
2.PL	lqat-t-u	t-o-lqt-u
3.PL	laqt-u	j-o-lqt-u

From the above paradigm we can see that the stem-form in the IMPERF sub-paradigm alternates between the SG vs. PL NUM values. Thus, to economise on the space available, it is only the 3SGM and 3PL values for the IMPERF paradigm that will be displayed. In the case of the PERF sub-paradigm, the 1SG cell will be taken to represent the underspecified 1<sup>2</sup> PERS. On the other hand, all three 3<sup>rd</sup> PERS slots will be displayed, even if in the above paradigm the 3SGF<sup>3</sup>PL stem-forms are syncretic. The need to render all three 3<sup>rd</sup> PERS cells is important as it is precisely on the basis of the distinct syncretic stem pattern organisation within these cells, that the PERF sub-paradigm of the distinct SBICs, varies. The paradigm for *laqat* 'hit' above, along with three others, to illustrate the variation in this ablaut-class, are given below.

ABCC SG≠PL <sup>7</sup>		laqat 'hit'	talab 'ask/pray'	hataf 'grasp'	harab 'escape'
		ABCC DE	ABCC DE	ABCC AD	ABCC AC
PERF	1.SG	lqat-t	tlab-t	htaft-t	hṛab-t
	3.SG.M	laqat	talab	hataf	harab
	3.SG.F	laqt-et	talb-et	hatf-et	harb-et
	3.PL	laqt-u	talb-u	hatf-u	harb-u
IMPERF	3.SG.M	j-o-lqot	j-i-tlob	j-a-htaf	j-a-hrab
	3.PL	j-o-lqt-u	j-i-tolb-u	j-a-htf-u	j-a-harb-u

An interesting variation can be seen across the different members of the a-a ablaut-class. Three distinct paradigmatic patterns are found, and are taken to represent sub-set of members within this ablaut class, which in turn transcends to having

<sup>7</sup> ABCC represents the 1<sup>2</sup>, 3SGM, and 3SGF<sup>3</sup>PL stem-indexing in the PERF sub-paradigm and SG ≠ PL implies that the IMPERF sub-paradigm is alternating. The SG and PL stem-slots in the IMPERF are not given a letter due to the fact that while this pattern of alternation is retained across the members of this SBIC, the stem-forms that are actually present in these stem-slots may not be new, but may involve a partial or total overlap of forms used in the PERF sub-paradigm, as shown to be the case in hataf 'grasp' and harab 'escape', respectively.

distinct sub-patterns within SBIC I.<sup>8</sup> The essential distinction between the *laqat-talab* and *hataf-harab* sets is that the former's pattern arises out of having five distinct stem-forms filling the five designated stem-slots in this class, implying that there are no overlapping forms across the sub-paradigms. In the latter set, on the other hand, while *hataf* has four stem-forms, which result from the syncretic stem-form across the PERF 1^2 and IMPERF SG realised by *h̄taf*, *harab* only has three, since its IMPERF sub-paradigm displays total overlapped forms. Just like the *hataf* class its PERF 1^2 and IMPERF SG stem-forms are syncretic (*h̄rab*-), however there is additional syncretism across the PERF 3SGF^3PL and IMPERF PL stem-slots (*h̄arb*-), which in turn, with both instances of syncretism, explain the three-fold stem-form alternation across the whole paradigm.

The significance of *talab* 'pray/ask' and *harab* 'escape', in relation to *laqat* 'hit' and *hataf* 'grasp' respectively, (where *laqat* and *talab* display the same pattern) is that, the stem's second consonant (C<sub>2</sub>) in *talab/harab* is a resonant, belong to the set of {m, n, l, r, gh} phonemes in Maltese. From this determining phonological similarity one would assume that these two verbs should behave in the same way in their respective paradigms, since this is the behaviour of *l* and *r* across the grammar, however they don't. Rather, as a result of something other than phonology, while the stem-vowel of the IMPERF SG stem-form in *talab* involves an ablaut change to *o* (*tlob*-) from the stem-vowel *a* of the PERF 1^2 stem-form (*tlab*-), *harab* does not involve this change, thus resulting in the stem pattern variation. Since *hataf-harab* do not involve ablaut change across the IMPERF SG and PERF 1^2 slots, retaining the stem-vowel *a* throughout, then, to account for the distinct number of stem-forms, there has to be stem-form overlapping across the lexemes' sub-paradigms, which further conflates features on the stem-form, resulting in stem-form losses. The loss of two stem-forms in the case of the *harab* set (vs. the one lost in the *hataf* set) comes about as a result of having a resonant C<sub>2</sub> coupled by a non-ablaut triggered stem-form change.

a-e

ABCC SG≠PL		ħasel 'wash'
		ABCC DE
PERF	1.SG	ħsil-t
	3.SG.M	ħasel
	3.SG.F	ħasl-et
	3.PL	ħasl-u
IMPERF	3.SG.M	j-a-ħsel
	3.PL	j-a-ħsl-u

e-e

ABCC SG≠PL		ħeles 'get rid'
		ABCC DE
PERF	1.SG	ħlis-t
	3.SG.M	ħeles
	3.SG.F	ħels-et
	3.PL	ħels-u
IMPERF	3.SG.M	j-e-ħles
	3.PL	j-e-ħils-u

<sup>8</sup> The reason for presenting the data as sub-patterns within distinct ablaut classes is to help the reader appreciate that the very same ablaut-membership need not imply sameness/homogeneity across classes traditionally considered to be the same.



These two ablaut-classes pattern exactly with the *laqat-talab* stem-pattern in the a-a ablaut class.

e-a

ABCC SG≠PL		bezaq ‘spit’ ABCC DE	feraq ‘split’ ABCC DE	mesah ‘wipe’ ABCC AD	ferah ‘be happy’ ABCC AD
PERF	1.SG	bzaq-t	fraq-t	msah-t	frah-t
	3.SG.M	bezaq	feraq	mesah	ferah
	3.SG.F	beżq-et	ferq-et	mesh-et	ferh-et
	3.PL	beżq-u	ferq-u	mesh-u	ferh-u
IMPERF	3.SG.M	j-o-bzoq	j-i/o-froq	j-i-msah	j-i-frah
	3.PL	j-o-bżq-u	j-i/o-forq-u	j-i-msh-u	j-i-firh-u

In the case of the e-a ablaut class, we have similar distinctions that match exactly with two of the three patterns we had in ablaut class a-a. The *bezaq-feraq* class patterns exactly like the *laqat-talab* class, with five distinct stem-forms filling in the five indexed stem-slots, whereas, interestingly, while the *harab* class differs from that of *hataf* in ablaut class a-a, *mesah-ferah* form one class, and pattern with the *hataf* sub-class having four stem-forms in their paradigms, with the forms *msah-* and *frah-* respectively being used across the PERF-IMPERF sub-paradigms in the PERF 3SGF^3PL and IMPERF PL stem-slots. Considering that the stem’s syllabic structure in the IMPERF PL cell of a resonant C<sub>2</sub> verb with a *strong* C<sub>[-w/j]</sub>VCVC verbal base such as *ferah* ‘be happy’ should be the same as that of the PERF 3SGF^3PL, i.e. CVCC, the fact that it still patterns in the same way as its non-resonant C<sub>2</sub> counterpart *mesah* ‘wipe’ implies that since *ferah* will not allow a \*CCC syllable-structured stem-form in the IMPERF PL cell, unlike *mesah*, than ablaut change is involved, which changes the stem-form from *ferh* to *firh*. If this were not the case, we would have expected the same stem-form overlap observed in the *harab* sub-class, rendering a syncretic stem-form in the PERF 3SGF^3PL and IMPERF PL slots.

i-e

ABCC SG≠PL		kiteb ‘write’ ABCC DE	tilef ‘lose’ ABCC DC
PERF	1.SG	ktib-t	tlif-t
	3.SG.M	kiteb	tilef
	3.SG.F	kitb-et	tilf-et
	3.PL	kitb-u	tilf-u
IMPERF	3.SG.M	j-i-kteb	j-i-tlef
	3.PL	j-i-ktb-u	j-i-tilf-u

In the i-e ablaut class, just like the *hataf-harab* pattern in ablaut class a-a, there is a varied stem pattern which cross-classifies across resonant and non-resonant C<sub>2</sub>

stems. While the latter class of verbs patterns with what seems to be the most common pattern in this SBIC, i.e. having five distinct stem-forms, the resonant C<sub>2</sub> class renders an interesting variation from the other patterns observed so far. What differs is that, unlike all other verbs with this characteristic resonant C<sub>2</sub> property in their stems, the IMPERF SG stem-form *tlef* does not overlap with that in the PERF 1^2 cell *tlif*, but rather includes a new distinct form. This class of verbs thus introduces a distinct sub-pattern that comes about as a result of having a distinct IMPERF SG stem-form and a syncretic one in the IMPERF PL slot.<sup>9</sup>

o-o

ABCC SG≠PL		xorob ‘drink’ ABCC AC
PERF	1.SG	xrob-t
	3.SG.M	xorob
	3.SG.F	xorb-ot
	3.PL	xorb-u
IMPERF	3.SG.M	j-i-xrob
	3.PL	j-i-xorb-u

This class of verbs follows exactly with all those verbs that pattern with *harab* in ablaut class a-a.

From the distinct patterns observed across one verbal base within SBIC I, it would be interesting to be able to come up with a parameter that could calibrate which of the stem patterns within a given SBIC appear to be the most canonical. To do so, in the meantime one would need to determine whether ablaut change, as opposed to syncretism, is preferred across the patterns. It seems intuitive to assume that if a SBIC stipulates a particular number of stem-slots, then the most canonical stem-pattern would be that which fills in these stem-slots with distinct forms. Hence, since from the paradigms above we can see that ablaut change renders distinct stem-forms, distinct stem-forms in every slot would be preferred over syncretic ones, as is the case in analogy with the different word-forms which fill a lexeme’s paradigmatic slots. It may seem that syncretism vs. ablaut change is correlated with ablaut class membership, where overlapping stem-forms have only been found in the a-a and o-o ablaut-classes. However, from the a-a ablaut class we see that this employs both ablaut changing and syncretic strategies, and thus, the correlation of ablaut class with a preference for a kind of stem-pattern over the other cannot quite be taken as an indicator of what one is likely to find.

---

<sup>9</sup> This class in general is interesting in the way ablaut changes seem to take place. While the IMPERF SG stem-form *tlef* seems to have retained the original 3SGM’s V2, this implies that the PERF 1^2 stem has involved e to i ablaut change, i.e. from a potential *tlef* to *tlif*, in this cell, instead of saying that the IMPERF SG stem involves an ablaut change from i to e (PERF 1^2 *tlif* to IMPERF SG *tlef*).

Before concluding this section, one should mention that while it was the verbal base  $C_{[-w/j]}VCVC$  discussed here, *weak-initial* verbal bases, whose initial C is a *w* (*assimilative*), including *wasal* ‘arrive’, *wiret* ‘inherit’, and *wizen* ‘weigh’, model exactly, and fit within the largest sub-pattern in this SBIC, having five distinct stem-forms. The purpose for this note here is to illustrate that it is surface form morphological patterns that account for SBIC memberships, and not phonological or UR accounts, as will also be displayed in the following two sections.

### 3.2.1.2 Non-CVCVC SBIC members

The paradigms discussed here should illustrate how SBIC membership is not determined by a type of morphophonological stipulation having to do with the verb’s morphophonological base. It is true however that one may say that, interestingly enough, the biggest majority of verbs in this class happens to have a CVCVC verbal base. Mixed memberships however, are what further make it even more of a morphologically-motivated system based solely on surface forms, without showing any relevance to the UR and with no reasonably explicable border with phonology, since SBIC membership need not necessarily cross-classify morphophonologically distinct verbal bases, as mentioned in section (3.1). While *mar* ‘go’ is the sole *hollow* verb that participates in this SBIC (see section 3.2.4 for the discussion on the class of traditional *hollow* verbs), this non-CVCVC verbal base membership is shared with the *a*-ablaut class of verbs belonging to the *lacking* verbal base (CVCV), illustrated below by *qara* ‘read’. While *mar* ‘go’ patterns with the majority stem pattern, interestingly, the *lacking* verb sub-set, which only includes  $C_2$  resonant stems, patterns exactly like the *tilef* class of verbs in ablaut class *i-e*, with a syncretic PERF 3SGF^3PL-IMPERF PL stem-form, and a distinct IMPERF SG form.

### 3.2.1.3 Irregulars

ABCC SG≠PL		mar ‘go’	ABCC SG≠PL		qara ‘read’
		ABCC DE			ABCC DC
PERF	1.SG	mor-t	PERF	1.SG	qraj-t
	3.SG.M	mār		3.SG.M	qara
	3.SG.F	marr-et		3.SG.F	qrā-t
	3.PL	marr-u		3.PL	qrā-w
IMPERF	3.SG.M	j-mūr	IMPERF	3.SG.M	j-a-qra
	3.PL	j-morr-u		3.PL	j-a-qrā-w

The sub-heading here may be confusing and will be subject to falsification. While the paradigm for *ha* ‘take’ below, in its surface pattern it is nothing but regular, and is a member of the largest sub-pattern in this SBIC, traditional Maltese pedagogical grammars treat this verb along with the other handful of *irregulars* mentioned in section (3.1). Its irregularity-based account stems from resorting to an UR, which however poses a stark contrast with what is observed from the surface paradigmatic pattern. It may only be suggestive of the fact that the verbal base inventory based on the

3SGM's stem-form should be expanded in SM, where we have to add a verbal base which includes only one C instead of the typical two or three. This is all that should be done in the overall grammar, instead of trying to impose that which is underlying onto the surface, precisely when affixation and stem-alternation patterns suggest an otherwise *non-irregular* verb. *Irregularity*, on this account, should only be taken to be something which comes out from the surface, as would be the case in an instance where there is a SBIC or stem pattern whose membership is highly restricted.

ABCC SG≠PL		ħa 'take' ABCC DE
PERF	1.SG	ħad-t
	3.SG.M	ħa
	3.SG.F	ħād-et
	3.PL	ħād-u
IMPERF	3.SG.M	j-ie-ħu
	3.PL	j-ie-hd-u

### 3.2.1.4 Summary

Our overview of SBIC I should have illustrated the subject-matter in further detail, focussing on the interesting array of variation within the same class, rendering four different sub-patterns. It was shown how a morphological pattern will not necessarily cross-classify the verbal bases of the morphophonological properties of the stem-form, or will not necessarily rely on morphotactic constraints, but is rather an autonomous morphological drive that classifies members, based on the stem-form organisation realising the syncretic stem index pattern.

### 3.2.2 SBIC II

This SBIC shows a distinct stem pattern which involves a non-alternating IMPERF sub-paradigm, and a pattern reorganisation in the PERF sub-paradigm, where the 3SGM stem-slot fuses with that of the 3SGF, while the 3PL takes its distinct stem-slot, as displayed in the ABBC vs. the ABCC pattern in SBIC I, illustrating how the PERF 3PL need not always be conflated with the 3SGF, justifying the reasoning behind the three-slotted 3<sup>rd</sup> PERS representation in the beginning of section (3.2.1.1). SBIC II is populated with the sub-class of *strong* verbs traditionally referred to as *deaf/geminate*, whose UR involves two identical Cs at the right-edge of the stem (or of the C-root), such as *ħass* 'feel'. Without needing to refer to any *underlying* morphological representations when referring to the members of this class, it suffices to say that the stem's C<sub>2</sub> and C<sub>3</sub> are the same. Since this class is exclusive to such verbs, it seems that we may have a case of a morphophonologically-conditioned SBIC membership, although see below.

ABBC SG=PL		ħass ‘feel’ ABBC DD → ABBB CC
PERF	1.SG	ħassej-t
	3.SG.M	ħass
	3.SG.F	ħass-et
	3.PL	<b>ħassē-w</b> ~ <b>ħass-u</b>
IMPERF	3.SG.M	j-ħoss
	3.PL	j-ħoss-u

As displayed from the above paradigm, the members of this class are *unstable*. *Unstability*, in our context here, should be understood through what can be illustrated from the above alternating pattern, which represents the whole class (except for verbs that pattern with *mess* ‘touch’ and *ghadd* ‘enumerate’, where these verbs participate in the ABBC SG=PL SBIC only, implying that their paradigms do not allow for an alternation in the PERF 3PL cell, and hence no shift in SBIC membership). This variation comes about as a result of the interaction with a further non-canonical morphologically-complex phenomenon; stem-form overabundance in the PERF 3PL cell, (illustrated in bold). This stem-overabundance, i.e. having more than one stem-form in a given cell, patterning the word-form overabundance which one gets when more than one word-form fills in a paradigmatic cell (Thornton 2010), creates a case of allomorphy for the inflectional material. While allomorphic alternation seems to be phonologically-conditioned, yet, interestingly, this has a further effect on the morphological features discussed here. Such allomorphy results in the possibility of having an additional complexity rendered by the availability of SBIC shifts. Thus, through the non-canonical occurrence of stem-overabundance we get what seems to be a non-canonical SBIC which allows its members to participate in two distinct SBICs. When the SBIC alternation takes place, the 3<sup>rd</sup> PERS stem-slots in the PERF sub-paradigm are levelled, resulting in a neat stem-split (Baerman/Corbett forthcoming) which renders a stem-alternation based on a discourse-participant vs. other PERS split (3<sup>rd</sup> vs. 1<sup>st</sup>), represented by the ABBB PERF sub-paradigm pattern. This process results in a reduction of a stem-slot and a stem-form, i.e. from four to three.

What should be said is that while a stem-form overabundance has here resulted in an SBIC-shift, this is not always necessarily the case, as we will see in section (3.2.4). At the same time, it is not only a non-canonical phenomenon such as overabundance that can cause this *non-canonical* SBIC shift/alternation in SM paradigms. Rather, this can also result from interactions external to the stem, such as when paradigms involve additional morphological elements, including the negation realising exponent, and attached pronouns. The other interesting thing which can be observed has to do with the relevance of the morphological island property which seems to motivate class formation. This is because when the *deaf/geminate* verbal base shifts its class membership, it starts patterning exactly in the same way as *hol-*

low verbs with a CV:C verbal base, as will be shown in section (3.2.4), which is taken to imply that SBIC membership is not triggered by any morphophonological stipulation from the verbal base. Although it appears that in many instances there is a cross-classification of SBIC membership and verbal base type, when this is not the case, as also illustrated in sections (3.2.1.2-3) one is able to appreciate and observe the independent status of *island morphology* at play.

**3.2.3 SBIC III**

SBIC III may be thought of as a heteroclitite class that takes the PERF stem pattern of SBIC II, with the indexed 1^2, 3SG, and 3PL stem-slots, and the more general alternating IMPERF SG≠PL pattern also found in SBIC I. This class thus displays the same number of stem-slots as SBIC I, however, since they are not the same class, then this means that their syncretic stem indexing pattern organises the slots differently, reflecting the distinct ways how features can be conflated together on the stem in the PERF sub-paradigm. This class illustrates what is here taken to be a real case of *irregularity*. While there is nothing *irregular* in the pattern as such, as the island morphological system is able to create whatever pattern the mechanism allows, what makes this pattern *irregular* in our understanding is that only one lexeme has this pattern in its paradigm. As a result of this single membership the class is regarded as non-canonical, by the criterion mentioned in section (2.2.1). What is further the case here with *gie* ‘come’ is that traditional Maltese grammars treat it along with *ħa* ‘take’, mentioned in section (3.2.1). Yet, these two verbs do not illustrate similar paradigmatic patterns. While the surface pattern of *ħa* is regular, belonging to the biggest SBIC, and can only be regarded as *irregular* if one is concerned with the verb’s UR, *gie*, which is also *irregular* in its UR, is here taken to be *irregular* only because it participates alone in its own separate class.

ABBC SG≠PL		gie ‘come’
		ABBC DE
PERF	1.SG	ġej-t
	3.SG.M	ġie
	3.SG.F	ġie-t
	3.PL	ġē-w
IMPERF	3.SG.M	j-i-ġi
	3.PL	j-i-ġ-u

**3.2.4 SBIC IV**

The pattern in this class has been mentioned in section (3.2.2) where members of SBIC II alternate with SBIC IV. This class involves a neat PERF-IMPERF distinction where unlike SBIC I there is no case of stem-form overlap across the two sub-paradigms, as shown in the representations from *zar* ‘visit’ and *far* ‘below’, with only one exception; *qal* ‘say’, to which we will return below. Furthermore, there is a neat

3 vs. 1^2 PERS split in the PERF sub-paradigm. Stem-Class IV is characterised by the traditional class of *hollow* verbs with a CV:C morphophonological verbal base, however, as shown from section (3.2.2) this base is not exclusive, and CVC<sub>i</sub>C<sub>i</sub> bases can also participate in this class. What is interesting to observe that while the class of verbs belonging to this SBIC is traditionally divided according to whether the verb has an underlying *-w-* or *-j-* medial consonant, from the surface form stem pattern representations this does not hold as a relevant distinction.

ABBB SG=PL		zar ‘visit’ ABBB CC	far ‘overflow’ <sup>10</sup> ABBB CC	ABBB SG=PL		qal ‘say’ ABBB CC
PERF	1.SG	zor-t	far/for-t	PERF	1.SG	ghid-t
	3.SG.M	zār	fār		3.SG.M	qāl
	3.SG.F	zār-et	fār-et		3.SG.F	qāl-et
	3.PL	zār-u	fār-u		3.PL	qāl-u
IMPERF	3.SG.M	j-zūr	j-fūr	IMPERF	3.SG.M	j-ghid
	3.PL	j-zūr-u	j-fūr-u		3.PL	j-ghid-u

The *zar-far* class of verbs, representing the whole of the SBIC members. *Qal* happens to represent its own separate way with which it realises the class’ pattern, which is not observed anywhere else across SM verbal paradigms. This pattern is essentially derived from the interaction of the non-canonical phenomena of suppletion and syncretism (rendering higher order exceptionality, Corbett 2011). Therefore, *qal* is not only *irregular* because of its particular stem pattern, but also because of the suppletion involved in the PERF sub-paradigm, which, coupled by the subsequent overlap of the suppleted form across the PERF-IMPERF paradigm, results in the only stem pattern that involves a PERF 1^2 stem-form used as the non-alternating stem of the IMPERF sub-paradigm, resulting in a stem-form loss. Unlike *gie* ‘come’, which has been treated as *irregular*, in the previous section, through evidence from its unique paradigmatic organisation of syncretic stem-slots, which happened to overlap with an irregular UR, *qal* is treated as *regular* from a UR perspective, at a par with other *-w-* medial verbs like *far* ‘overflow’ in traditional grammars. In our study here we see that *qal* displays an *irregular* means with which it realises its stem index pattern. For this reason, classifying it with the rest of the *hollow* class in which it belongs, misses out on the interesting variation and the contribution this paradigm gives to the analysis of the realisation of SBICs across 1<sup>st</sup> *binyan* SM verbal paradigms. The same follows for *mar* ‘go’, which does not belong in this SBIC altogether, as mentioned in section (3.2.1.2). If we look

<sup>10</sup> *Far* ‘overflow’ illustrates a case of stem-allomorphy, which, contra to the stem-allomorphy in the *hass* ‘feel’ paradigm, this does not result in a SBIC shift. While in SBIC II it was a change in the stem’s syllable-structure that rendered the change, this cannot always be taken to be an indicator of potential SBIC shifts. This fact is illustrated by a sub-set of lexemes in this class, such as *gham* ‘swim’, which involve a syllable-structure based stem-alternation: *ghom* ~ *ghomej* in the PERF 1^2 stem-slot, where like *far*, does not involve a shift in SBIC membership.

at the alternating paradigms below represented here by *sam* ‘fast’, we observe that we get a case of overabundance in the PERF 3PL cell, and not in the PERF 1^2 cell, as is the case with *far* ‘overflow’, *gham* ‘swim’, mentioned in (ftn. 9), and other members of this class. The alternation here parallels that which we had in SBIC II, illustrated by *hass* ‘feel’, even though the PERF 1^2 alternation is also possible here, which as also mentioned in (ftn. 10) above, will not lead to any distinct results. While there it was a case of a SBIC II to SBIC IV alternation, here we have the reversed direction i.e. an alternation from SBIC IV to SBIC II, which results in an additional stem-form. What one wants to say here is that the overabundant alternation comes out from a redundant morphological selection of a distinct allomorph realising the PL features in this cell. As a result of this redundant change in form, the phonology comes to ensure that the stem-form’s form fits the phonological requirements.

ABBB SG=PL		sam ‘fast’ ABBB CC → ABBC DD
PERF	1.SG	som-t
	3.SG.M	sām
	3.SG.F	sām-et
	3.PL	<b>sām</b> -u ~ <b>samē</b> -w
IMPERF	3.SG.M	j-sūm
	3.PL	j-sūm-u

### 3.2.5 SBIC V

SBIC V includes two members, each with its own separate stem pattern, and just like SBIC III appears to be heteroclite, involving a PERF sub-paradigm that patterns the stem-pattern organisation in SBIC IV having a 3 vs. 1^2 stem-alternation, and an alternating IMPERF sub-paradigm that patterns with SBIC I and III. What is interesting to observe is that the *irregular* heteroclite SBICs III and V overlap with the majority stem-slot organisatory pattern in SBIC I, in their IMPERF pattern. Just like *gie* ‘come’ in stem-class III, *kiel* ‘eat’ and *ra* ‘see’ are also treated as morphologically *irregular* in traditional grammar, on the basis of the odd and obsolete C-root in their UR. Although this SBIC is made up of two members, and may be regarded as *less irregular* than SBIC III, which only has one member, yet, the two members have their own separate stem patterns, which individuates them once again. While *kiel* ‘eat’ matches its number of stem-slots with the same number of stem-forms, *ra* ‘see’ involves an overlapping stem-form, thus having three stem-forms filling in the four stem-slots.



ABBB SG≠PL		kiel ‘eat’ ABBB CD
PERF	1.SG	kil-t
	3.SG.M	kiel
	3.SG.F	kiel-et
	3.PL	kiel-u
IMPERF	3.SG.M	j-ie-kol
	3.PL	j-ie-kl-u

ABBB SG≠PL		ra ‘see’ <sup>11</sup> ABBB CB
PERF	1.SG	raj-t
	3.SG.M	ra
	3.SG.F	ra-t
	3.PL	ra-w
IMPERF	3.SG.M	j-à-ra
	3.PL	j-a-rà-w

### 3.2.6 SBIC VI

SBIC VI is characterised by having the highest number of stem-slots in its members’ paradigms, whose syncretic stem organisation pattern renders six distinct stem-slots in the overall paradigm. While the IMPERF’s pattern is not innovative and alternates on the basis of an SG vs. PL distinction, in the PERF sub-paradigm we have four stem-slots: 1<sup>^</sup>2, 3SGM, 3SGF, 3PL, which pattern is thus the main motivator behind having the four distinct PERF stem-slots represented in our data. As we will see, this class is divided in two stem patterns; one that has six distinct stem-forms, each one mapping to a designated stem-slot, and the other having five forms, involving an overlapping stem-form across the sub-paradigms.

The members of this SBIC are all members of one of the two ablaut classes found in the traditional class of *lacking* verbs; e-a. The other class of verbs belonging to the a-a ablaut class participate in SBIC I, as mentioned in section (3.2.1.2). Interestingly then, the class of CVCV verbal bases is divided into separate SBICs on the basis of the ablaut class they belong to, which is not what happens across the different ablaut classes of the traditional *strong* class, which, as we have seen in section (3.2.1.1), all participate in the same class. It is important to mention that traditional Maltese grammars only distinguish different members of the *lacking* class on the basis of whether the final weak consonant (which never surfaces in verbal paradigms) is -w or -j. The a-a/e-a ablaut class distinction is not given any independent attention as something that characterises the *lacking* class, in contrast with the great deal of attention given to the fact that there exist six ablaut classes in the *strong* class. The importance given to the *lacking* class’ ablaut class division only surfaces to illustrate the significance these distinct ablaut classes had within the diachronic process involved in the integration of the Romance part of the language within the set Semitic verbal paradigmatic model, and are thus only attended to in relation with diachrony (Mifsud, 1995, Hoberman/Aronoff, 2003). What is being proposed here is that, on the contrary, these two ablaut classes are significant within the synchronic context. Furthermore I go a step ahead to say that, synchronically, due to what we observe on the basis of paradigmatic data, the distinction across the a-a vs. e-a ablaut classes is

<sup>11</sup> This is treated as a separate stem pattern that includes alternation in the IMPERF sub-paradigm on the basis of stress-variation across the SG-PL stem-slots.

more relevant than that found in the *strong* class, as the distinction in the *lacking* class has been grammaticalised and rendered as a split membership in two SBICs.

ABCD SG≠PL		heba ‘hide’ ABCD EF	kera ‘hire’ <sup>12</sup> ABCD EF	beda ‘start’ ABCD ED	mela ‘fill’ ABCD ED
PERF	1.SG	hbej-t	krej-t	bdej-t	mlej-t
	3.SG.M	heba	kera	beda	mela
	3.SG.F	hbie-t	krie-t	bdie-t	mlie-t
	3.PL	hbē-w	krē-w	bdē-w	mlē-w
IMPERF	3.SG.M	j-a-hbi	j-i-kri	j-i-bda	j-i-mla
	3.PL	j-a-hb-u	j-i-kr-u	j-i-bdē-w	j-i-mlē-w

The above paradigms illustrate how ablaut class e-a is further divided into two distinct sub-classes that essentially depend on the nature of the stem-vowel present in the IMPERF sub-paradigm. While in the *heba-kera* class there is an ablaut change from PERF 1^2 *e* (*hbej*) to IMPERF SG *i* (*hbi*), in the *beda-mela* class the PERF 1^2 *e* changes to IMPERF SG *a* (*bdej* to *bda*), which is the pattern that involves a syncretic form across the PERF 3PL and IMPERF PL. It is important to emphasise on the fact that this ablaut change does by no means cross-classify any UR, rather, it happens that all the verbs in the above four paradigms have a *-j* final radical, and where the *heba-kera* sub-class includes both *-j* and *-w* final verbs. Precisely because there is no logically possible explanation as to why the *beda-mela* class does not function in the same way as the *heba-kera* class or vice-versa, all this is taken to be significant as it implies that SBICs and stem-patterns do really involve internal drives, and are not affected by the external syntax or phonology.

#### 4. CONCLUSION

From the above 1<sup>st</sup> *binyan* verbal data from SM, we can now define the *stem-based inflectional class* as a morphological phenomenon that arises as a result of the non-adherence of internal-paradigmatic canonicity, which follows from the treatment of the stem as pure lexical material that requires to be inert and non-alternating. SBICs are thus formed as a result of different indexing patterns of syncretic stem-slots, whose pattern of stem-form realisation within the same organisatory pattern may vary, and has to do with the number of stem-forms filling the available slots. As a result of alternating stems in paradigms, the canonical distinction/split between affixes as inflectional material vs. stems as lexical material, is blurred, which in SM verbal paradigms results in the conflation of morphological features on the stem, and/or feature overlap

<sup>12</sup> Note that although in the paradigms represented by *kera* ‘hire’ and *mela* ‘fill’ there is a resonant C<sub>2</sub> in the stem, these however do not behave in the same way as the resonant C<sub>2</sub> in SBIC I stems. This difference is because while in SBIC I we had stems with three Cs and the resonant C<sub>2</sub> happened to be in a medial position, thus causing the effects observed in section (3.2.1.1), in SBIC VI, the stems only have two consonants, and thus, the resonant C is now in a final position.

with the affixal material, both resulting in multiple exponence (Baerman/Corbett 2010). Having established what such classes look like in SM, on the basis of the data in section (3.2), we have seen that there are six SBICs that classify the 1<sup>st</sup> *binyan* verbs. While SM appears to allow up to six distinct stem-slots across the PERF-IMPERF paradigm, it is only SBIC VI that fulfills this maximal distinction, and is only restricted to a small sub-set of lexemes, which seems to have resulted from an actual grammaticalisation of something that was merely a phonological ablaut class difference. It was further shown how with the classification of SM verbs in such classes, surface form morphology may be in a mismatched relation with the traditional UR of these same verbs. It is not just the UR that does not necessarily predict SBIC membership and organisation, but also not even surface structure morphophonological similarities. This was shown to be the case especially when the same verbal base participates in more than one SBIC or when distinct verbal bases participate in the same SBIC. With these occurrences, we can clearly bring out the internal autonomous drive upheld by the morphological component, especially when its forces override the potential border-bridging and interaction with the phonological interface, hence reinstating further morphology's *island*-status.

## References

- ALEXIADOU, Artemis (2004) "Inflection class, gender, and DP internal structure." In: G. Müller (ed) *Exploration in Nominal Inflection*. Berlin: Walter de Gruyter, 21-49.
- AQUILINA, Joseph (1973) *The Structure of Maltese*. University of Malta, Malta.
- BAERMAN, Matthew (2011) "Inflectional interactions in Oto-Manguaen." (January Conference).
- BAERMAN, Matthew/BROWN, Dunstan/CORBETT, Greville G. (2005) *The Syntax-Morphology Interface: A study of syncretism*. Cambridge: CUP.
- BAERMAN, Matthew/CORBETT, Greville G. (forthcoming) "Stem alternations and multiple exponence." Paper presented at the workshop 'Stems in Inflection and Words'.
- BROWN, Dunstan. (2008) "Morphological Typology." In: J. J. Song (ed) *The Oxford Handbook of Linguistic Typology*. Oxford: OUP, 487-503.
- CORBETT, Greville G. (2005) "The canonical approach in typology." In: Z. Frajzyngier/D. Rood/A. Hodges (eds) *Linguistic Diversity and Language Themes*. Amsterdam: Benjamins, 25-49.
- CORBETT, Greville G. (2007) "Canonical typology, suppletion, and possible words." *Language* 83, 8-42.
- CORBETT, Greville G. (2009) "Canonical Inflectional Classes." In: F. Montermini/G. Boyé/J. Tseng (eds) *Selected proceedings of the 6th Decembrettes*. Somerville: Cascadilla Proceedings Projects, 1-11.
- CORBETT, Greville G. (2011) "The prenumbra of morphosyntactic feature systems" *Morphology* 21, 445-480.
- CORBETT, Greville G. (2011) "Higher order exceptionality in inflectional morphology." In: H. J. Simon/H. Wiese (eds) *Expecting the unexpected: Exceptions in grammar* (Trends in Linguistics. Studies and Monographs 216), Berlin: De Gruyter Mouton, 107-126 [With

- comments by Stephen Anderson, “An I-language view of morphological ‘exceptionality’: comments on Corbett’s paper”, 127-134, and reply “Exceptions and what they tell us: reflections on Anderson’s comments”, 135-137].
- CORBETT, Greville G./BAERMAN, Matthew (2006) “Prolegomena to a typology of morphological features” *Morphology* 16, 231-246.
- CORBETT, Greville G./BAERMAN, Matthew (2010) “Deep Irregularity”: When the expected lexical splits follow the wrong pattern. Paper presented at the 14th International Morphology Meeting, Budapest, Hungary.
- HOBERMAN, Robert D./ARONOFF, Mark (2003) “The verbal morphology of Maltese: From Semitic to Romance.” In: J. Shimron (ed) *Language Processing and Acquisition in Languages of Semitic, Root-based, Morphology*. Amsterdam: Benjamins, 61-78.
- MIFSUD, Manwel (1995) *Loan verbs in Maltese*. Brill: Leiden.
- RYDING, Karen C. (2005) *A Reference Grammar of Modern Standard Arabic*. Cambridge: CUP.
- THORNTON, Anna M. (2010) *Overabundance, shape conditions and paradigms*. University of Surrey.

#### Abstract

### ISLAND MORPHOLOGY: MORPHOLOGY’S INTERACTIONS IN THE STUDY OF STEM PATTERNS

The paper discusses the notion of morphological complexity, with a focus on stem patterns. Stem patterns, creating stem-based inflectional classes, are morphological constructs which come about as a result of observing the patterns rendered by the stem-form alternations (or stem splits (Baerman/Corbett forthcoming)), which one extracts after the formation of word-forms within paradigms. Stem-based inflectional class formation constitutes one aspect in the analysis of non-canonical paradigms, which also include affix-based inflectional classes, syncretism, defectiveness, and overabundance Corbett 2005, 2007, 2009; Baerman, Brown and Corbett 2005; Thornton 2010). While these non-canonical instances are in themselves interesting to observe, it is even more intriguing to be able to see what interactions can arise, which at times do not seem to be the result of something exterior to morphology proper. Through data taken from Maltese verbal paradigms the phenomenon of stem-based inflectional classes will be explored, which will exhibit how internal to the paradigm there exists a complex system in itself, which is based on the distinct organisation of different conflated morphosyntactic features which come about via syncretism. These patterns should illustrate a paradigm-internal morphological phenomenon that is irrelevant to the syntax, where while morphology borders with it, there need not be any interaction at this interface. At the same time, it will also be shown how at times, the border with phonology is blurred, where while the phonology may often try to build bridges that interface with the morphological island, the island’s internal forces that drive its autonomy may deem to be more superior than the phonology’s strive to impose its interacting requirements, which render some interesting morphophonological mismatches as a result.

Povzetek

OTOŠKA MORFOLOGIJA: VLOGA MORFOLOŠKIH INTERAKCIJ  
PRI RAZISKOVANJU VZORCEV OSNOVE

Članek se posveča pojmu morfološke kompleksnosti s posebnim poudarkom na vzorcih osnove. Vzorci osnove, na katerih se oblikujejo na osnovah temelječi pregibni razredi, so morfološki konstrukti, ki nastanejo kot rezultat upoštevanja vzorcev sprememb oblike osnove (oziroma razcepov osnove /ang. *stem splits*/ (Baerman/Corbett v tisku)), kot jih lahko ugotovimo na podlagi tvorjenja besednih oblik znotraj posamezne paradigme. Tvorjenje, ki zajema na osnovah temelječe razrede, predstavlja enega od vidikov analize nepravilnih paradigem, ki vključujejo tudi pregibne razrede, sinkretizem, nezapolnjenost in prenapolnjenost (Corbett 2005, 2007, 2009; Baerman, Brown and Corbett 2005; Thornton 2010). Proučevanje teh nepravilnih primerov je že samo na sebi zanimivo, a večji izziv predstavlja ugotavljanje, do kakšnih medsebojnih vplivanj lahko pride, za katera pa se včasih zdi, da ne izvirajo iz nečesa, kar je zunaj morfološke v ožjem pomenu besede. Na podlagi gradiva, ki ga tvorijo malteške glagolske paradigme, bomo raziskovali pojav na osnovah temelječih pregibnih razredov in poskušali ugotoviti, kako kompleksen sistem obstaja sam na sebi znotraj vsake posamezne paradigme, sistem, ki izvira iz neodvisnega ustroja različnih morfosintaktičnih lastnosti, ki so posledica sinkretizma. Ti vzorci naj bi bili primer morfološkega pojava, ki je omejen na posamezno paradigmo in ki ne sega na področje sintakse, saj naj, četudi morfološka z njo meji, med njima ne bi prihajalo do medsebojnih vplivanj. Istočasno bomo pokazali, kako se, čeprav si fonologija pogosto prizadeva zgraditi mostove z morfološkim otokom, občasno zabrišejo meje s fonologijo, saj se lahko notranjim otoškimi silami, ki so gonilo njegove avtonomije, pripiše vpliv, ki presega fonološki napor, ki želi vsiliti svoje težnje po vplivanju, od koder izvirajo nekatera zanimiva morfofonološka neskladja.



## UN CAS D'EXTRAGRAMMATICALITÉ PARTICULIER : LES AMALGAMES LEXICAUX FANTAISISTES

### INTRODUCTION

Si Sablayrolles (2000 : 224) dénombre une dizaine de termes tels que mot-valise, mot-centaure, mot-gigogne ou mot-tiroir pour désigner les amalgames lexicaux, ces unités n'ont paradoxalement pas beaucoup bénéficié du regard des linguistes. Seules quelques études – qui seront détaillées *infra* – concernent la place de ces unités lexicales en français. Plusieurs facteurs sont susceptibles d'expliquer ce manque d'intérêt. Galisson (1987 : 58) ou Sablayrolles (1997 : 30) soulignent notamment leur caractère éphémère. La collecte de données s'en trouve alors entravée bien que l'utilisation d'outils informatiques<sup>1</sup> confère plus de facilité dans cette tâche<sup>2</sup>.

L'amalgamation lexicale mérite pourtant que l'on s'y intéresse, peut-être justement parce qu'elle relève de l'extragrammatical. En effet, cette matrice de création lexicale est particulièrement productive<sup>3</sup>. Des amalgames lexicaux sont apparus dans des lexiques de spécialité comme la publicité, le commerce ou encore la physique et la chimie. L'intérêt didactique<sup>4</sup> de l'amalgamation lexicale est également un élément à prendre en considération.

Dans cet article, nous ne reviendrons pas sur la nature extragrammaticale des amalgames lexicaux puisque nous adhérons à l'étude conduite par Fradin, Montermini et Plénat (2009). En revanche, nous désirons dresser un état des lieux des recherches actuelles et initier une description pour mieux cerner le phénomène. Cette description s'appuie sur un corpus original constitué de 7052 amalgames lexicaux issus de dictionnaires fantaisistes. Ce corpus, de nature similaire à celui de Galisson (1987) mais de taille bien plus importante, doit fournir autant d'informations sur les structures morpho-phonologiques de ces unités que possible, dans le cadre d'une création lexicale *in vitro*. Les résultats obtenus sont ensuite confrontés à un second corpus composé d'amalgames lexicalisés, voire lexicographiquement attestés, issus de différentes sources. Ce second corpus est, de fait, quantitativement moins important que le premier, dans la mesure où nous avons déjà pointé la difficulté à rassembler des données attestées. La mise en regard de ces deux corpus per-

---

\* Adresse de l'auteur: LDI, Université de Cergy-Pontoise, 33 boulevard du Port, 95011 Cergy-Pontoise, France. Mél: arnaud.leturgie@gmail.com

<sup>1</sup> La base de données néologique Neologia du laboratoire LDI (CNRS/universités Paris 13 et Cergy-Pontoise) fournit un nombre important de données pour la collecte d'amalgames lexicaux.

<sup>2</sup> Cependant, il n'existe pas, à notre connaissance, un outil permettant l'identification automatique des amalgames lexicaux.

<sup>3</sup> Fradin, Montermini et Plénat (2009).

<sup>4</sup> Cf. Galisson (1987) et Léturgie (à paraître) sur ce sujet.

met de mettre en lumière des correspondances et des divergences qui offrent une description générale de l'amalgamation lexicale.

Dans un premier temps les principales approches théoriques de l'amalgamation lexicale seront présentées. Cela sera l'occasion d'évaluer les différences qui les séparent et les lacunes qu'elles peuvent présenter. Puis, une approche prototypiste du phénomène sera introduite. Cette approche permet notamment de s'affranchir des divergences qui divisent les études antérieures afin d'offrir un cadre de recherche plus souple. Dans un second temps, ce sont les corpus sur lesquels nous nous appuyons qui seront présentés. La mise en relation d'un corpus *in vitro* et d'un corpus *in vivo* illustre un rapport similaire à celui entretenu par les morphologies grammaticale et extragrammaticale. Enfin, la recherche d'un prototype pour le français sera amorcée. La description proposée dans cet article doit permettre de mieux percevoir l'amalgamation de façon générale, notamment pour apporter de la documentation sur l'extragrammaticalité du phénomène.

## **1. L'AMALGAMATION LEXICALE : DIFFÉRENTES APPROCHES THÉORIQUES**

Comme annoncé en introduction, plusieurs études ont été menées sur l'amalgamation lexicale en français. Mais ces études sont assez anciennes (milieu des années 80) et n'ont pas fait l'objet de recherches suivies. De plus, les méthodes d'identification et d'analyse des amalgames lexicaux proposées dans ces études diffèrent et ne reposent pas sur les mêmes critères. Les quatre approches les plus significatives pour le français présentent effectivement des différences significatives qu'il faut souligner. Les trois premières sont apparues en trois années d'intervalle et suivent des méthodes bien différentes. Alors que la première approche de Grésillon (1984) est, encore aujourd'hui, l'étude centrale du phénomène pour le français, celles de Clas (1987) et Galisson (1987) sont moins souvent citées, particulièrement pour la dernière. Plus récemment, Fradin (1997 et 2000) et Fradin, Montermini et Plénat (2009) ont évoqué les aspects sémantiques et phonologiques des amalgames lexicaux en français.

Précisons d'emblée que l'aspect sémantique de l'amalgamation lexicale ne sera pas traité dans cet article. Les corpus sur lesquels sont basées les recherches présentées ici n'ont pas encore fait l'objet d'un dépouillement de type sémantique. Afin de livrer une étude aussi précise que possible, nous limitons volontairement notre approche à la morphologie et à la phonologie.

Ces quatre approches vont être présentées ici. Puis, un modèle théorique assez récent proposant de procéder par prototypie, et que nous avons décidé d'adopter, sera introduit.

### **1.1. Approches pour le français**

Puisque l'étude que nous proposons concerne le français, seules les approches concernant cette langue seront développées ici. Les quatre approches retenues sont les plus significatives mais ne proposent pas nécessairement de situer la place de



l'amalgamation lexicale au sein de la morphologie. En effet, Clas (1987) et Galisson (1987) traitent le phénomène sans insister sur son caractère extragrammatical alors que Grésillon (1984) qualifie de « monstres » les amalgames lexicaux pour traduire ce caractère. Enfin, Fradin (2000) puis Fradin, Montermini et Plénat (2009) installent l'amalgamation lexicale au sein de l'opposition entre morphologie grammaticale et morphologie extragrammaticale.

### **1.1.1. Une approche globale : Grésillon (1984)**

La première étude d'importance est celle conduite par Grésillon (1984). C'est encore aujourd'hui l'étude de référence pour les chercheurs qui s'intéressent au phénomène de l'amalgamation lexicale en français. Elle détermine des critères définitoires permettant d'identifier et de catégoriser les amalgames lexicaux. Selon elle, une unité lexicale doit être dotée de plusieurs propriétés (p. 12-13) pour être considérée comme un amalgame. Cependant, c'est la présence d'un segment homophone, ou quasi-homophone aux bases-sources qui incarne « la condition *sine qua non* de la formation du mot-valise » (p. 15). Elle envisage donc quatre schémas de formation (p. 24-25) :

- avec segment homophone = *enfantaisie* < enfant + fantaisie ;
- avec troncation = *infiniverti* < infini + inverti ;
- avec segment homophone et troncation = *éphémérde* < éphémère + merde ;
- avec enchâssement = *enseignement* < enseignement + saigne.

Cette approche a l'avantage de décrire la formation des amalgames d'un point de vue morpho-phonologique et les schémas proposés s'avèrent assez complets pour être appliqués à des corpus de grande ampleur. Pourtant, le corpus principal sur lequel repose ce travail est composé d'une soixantaine de termes allemands créés par Heine, poète allemand du XIX<sup>e</sup> siècle, complété d'un corpus annexe de nature très hétérogène<sup>5</sup>. Or, étant donné la difficulté de tirer beaucoup de données à partir du corpus réduit sur lequel elle s'appuie, ces schémas semblent manquer de détails. Mais si Grésillon ne prend pas la peine de détailler les combinaisons morphologiques des amalgames qu'elle étudie, c'est peut-être que ces combinaisons étant imprédictibles – ce qui constitue d'ailleurs une des propriétés qui fait des amalgames lexicaux des unités morphologiquement extragrammaticales – il suffit de mentionner le phénomène de troncation. Néanmoins, situer le lieu de la troncation dans les bases-sources ne semble pas superflu. L'aspect phonologique est mieux documenté puisque les schémas de recouvrements homophoniques sont présentés (p. 19-20).

Un des problèmes à soulever est que les critères semblent devoir convenir aussi bien pour le français que pour l'allemand et l'anglais. Et on imagine qu'il doit en

---

<sup>5</sup> Le corpus de Grésillon (1984) est constitué de termes issus de différentes sources telles que le journalisme, le commerce, la publicité ou l'édition, mais aussi de termes forgés particulièrement pour l'étude, ainsi que de répertoires de mots fantaisistes. Trois langues (allemand, français et anglais) sont utilisées et traitées indifféremment.

être de même pour toutes les autres langues. Seulement, des propriétés importantes – qui ne sont pas toujours partagées – de ces langues sont à prendre en compte, comme la structure accentuelle des unités lexicales. L'accentuation va jouer un rôle considérable dans l'opération de segmentation puis d'amalgamation des bases-sources dans des langues comme l'anglais<sup>6</sup> ou l'hébreu<sup>7</sup>.

### **1.1.2. Typologie morphologique : Clas (1987)**

Clas (1987) s'appuie sur un corpus de termes techniques, essentiellement issus de la chimie ou de la physique. Son analyse repose totalement sur la segmentation morphologique des unités lexicales. Il dresse donc une typologie des amalgames lexicaux en considérant 6 schémas de troncation. C'est effectivement selon la place où intervient la troncation que l'amalgame lexical est catégorisé. Il peut donc s'agir d'unités formées par:

- aphérèse d'une base-source = *robotique* < robot + informatique ;
- apocope d'une base-source = *publipostage* < publicité + postage ;
- aphérèse de la première base-source et apocope de la seconde base-source = *domotique* < domicile + robotique ;
- aphérèse des deux bases-sources = *nylon* < vinyl + coton ;
- apocope des deux bases-sources = *télex* < teleprinter + exchange ;
- aphérèse ou apocope d'une base-source et syncope de la deuxième base-source = *upérisation* < ultra + pasteurisation.

Bien qu'il soit morphologiquement plus détaillé que celui de Grésillon (1984), le modèle de Clas (1987) ne prend pas en considération tout l'aspect phonologique. Les cas d'homophonie indiqués par Grésillon ne peuvent être identifiés par l'approche de Clas alors que ces cas paraissent représenter plus fidèlement le concept d'amalgame lexical. Clas n'évoque finalement que la dimension concernée par l'ex-tragrammaticalité. C'est en effet la morphologie, dégagée de ses interactions avec la phonologie, la syntaxe ou la sémantique qui est envisagée dans sa typologie. Le problème est que, ce faisant, le chercheur ne peut pas livrer tous les schémas nécessaires à une description de l'amalgamation lexicale. Nous verrons au §3 que ce modèle doit être adapté pour devenir réellement opératoire.

### **1.1.3. Approche ludo-didactique : Galisson (1987)**

L'étude proposée par Galisson (1987) est assez atypique<sup>8</sup>. En effet, le corpus sur lequel il s'appuie est, à l'image de celui que nous proposons aujourd'hui<sup>9</sup>, composé entièrement d'amalgames fantaisistes puisqu'il repose sur 3 répertoires fantaisistes,

---

<sup>6</sup> Bertinetto (2001) ou Gries (2004).

<sup>7</sup> Bat-El (1996).

<sup>8</sup> Cf. Léturgie (à paraître) pour plus de détails.

<sup>9</sup> Cependant, le traitement qu'il applique à ce corpus est très différent de notre méthodologie.

dont son propre *Distractionnaire* (1986), ce qui représente environ un millier d'unités. Galisson distingue les unités qui relèvent de l'insertion de celles qui relèvent de l'imbrication, ces deux matrices connaissant des variations :

- insertion sans changement graphique = *ministère* < mini + stère + ministère ;
- insertion avec changement graphique du mot originel = *mélancolis* < mélancolie + colis ;
- insertion avec altération du mot originel = *néofrite* < néophyte + frite ;
- insertion avec substitution d'une partie du mot originel = *merfidie* < mère + perfidie ;
- imbrication = *ringargariser* < ringard + gargariser.

Les schémas proposés semblent être assez souples pour s'adapter à un corpus de grande ampleur. Cependant, la description du phénomène demeure assez pauvre et il n'est fait aucune mention de l'aspect morpho-phonologique de l'amalgamation lexicale. La typologie résultante est donc moyennement satisfaisante et l'approche préconisée par Grésillon (1984) semble encore la plus indiquée.

#### **1.1.4. Les études récentes de Fradin**

Enfin, plus récemment, Fradin (1997) et Fradin, Montermini et Plénat (2009) ont également tenté de décrire le phénomène. Fradin (1997) centre son analyse sur la sémantique des amalgames lexicaux. Fradin (2000) et Fradin et *al.* (2009) laissent plus de place à la description morpho-phonologique de ces unités lexicales.

Fradin (2000) et Fradin et *al.* (2009) offrent une description assez précise du phénomène en s'appuyant sur des données antérieures assez anciennes (dont Grésillon 1984). L'agencement phonologique est particulièrement observé, notamment en ce qui concerne les cas d'homophonie des bases-sources. L'analyse proposée par Fradin (2000 : 28-30) et Fradin et *al.* (2009) est très aboutie et a l'avantage de reposer sur de nombreuses études antérieures. Les schémas de construction phonologique sont très nombreux et détaillés, à tel point que nous ne disposons pas de la place nécessaire pour les reproduire ici.

Ces différentes approches ne pointent pas toutes dans la même direction. Ainsi, afin de ne pas se limiter à suivre telle ou telle approche, Renner (2006) préconise une description de l'amalgamation lexicale par prototypie.

## **1.2. Approche prototypiste**

Les divergences théoriques qui viennent d'être illustrées par ces quatre approches ne concernent pas que la recherche francophone. Brdar-Szabó/Brdar (2008) notent également une absence de consensus entre les chercheurs.

There are so many things about [lexical blending] that remain to be investigated and so many issues on which there is no real consensus among researchers, from a satisfying definition to an extensional account of the phenomenon including an inventory of various subtypes. Brdar-Szabó/Brdar (2008 : 172)

Renner (2006) remarque lui aussi une multiplicité de points de vue sur le phénomène. Il propose néanmoins d'utiliser ces différentes approches afin de dégager une définition prototypique de l'amalgamation lexicale. En synthétisant les quatre contraintes les plus fréquentes parmi les études réunies, il dévoile un prototype général d'amalgame lexical. Cette approche est tout particulièrement intéressante dans la mesure où elle confronte non seulement des points de vues sur un même phénomène mais qu'elle s'inscrit dans une perspective contrastive puisqu'elle repose sur des études en français, anglais, espagnol, allemand et bulgare.

Il s'agira donc ici<sup>10</sup> de s'appuyer sur cette étude afin d'illustrer les différentes façons d'appréhender ce phénomène. À partir d'une définition dite du « plus petit dénominateur commun » commune à toutes les théories, plusieurs critères défini- toires sont exposés afin de témoigner des divergences entre les différentes approches. Trois critères partagent les spécialistes sur la question et nous allons les évoquer brièvement ici à titre de rappel.

### **1.2.1. « Le plus petit dénominateur commun »**

Revenons en tout premier lieu sur ce que Renner appelle « PPDC »<sup>11</sup> : la définition minimale implicitement acceptée par les linguistes lorsqu'il s'agit de définir l'amalgamation lexicale. Ainsi, « l'amalgame est issu de l'assemblage d'au moins deux lexèmes, processus au cours duquel au moins un des lexèmes perd une partie de son signifiant »<sup>12</sup>. C'est une définition qui est, comme le souligne Renner, suffisante pour de nombreux linguistes<sup>13</sup> ayant traité la question. Cependant, elle est insuffisante selon d'autres chercheurs et d'autres critères entrent en jeu pour déterminer de l'appartenance ou non à la catégorie des amalgames lexicaux.

### **1.2.2. Contrainte morphologique : TRONC INT**

La première contrainte évoquée par Renner (2006 : 139) est la troncation interne (TRONC INT). Il s'agit de la caractéristique la plus courante et elle est d'ailleurs adoptée par de nombreux linguistes<sup>14</sup>. Cette contrainte « veut que le schéma de troncation corresponde à une apocope de la première base-source et/ou une aphérèse de la deuxième base-source » (Renner, *ibid.*). En d'autres termes, les mots composés par deux bases-sources apocopées ou aphérésées comme *chloroforme* (< chlore + for- myle) ; *pixel* (< picture + element) ou *nylon* (< vinyl + cotton) n'appartiendraient pas à la catégorie des amalgames lexicaux<sup>15</sup>. Cette contrainte en appelle une autre : la

---

<sup>10</sup> Cette partie est identique à ce qui est présenté dans Leturgie (2011 : 77-79).

<sup>11</sup> Pour « plus petit dénominateur commun » (Renner 2006 : 139).

<sup>12</sup> *Ibid.*

<sup>13</sup> Sont cités Adams (1973 : 137, 146), Algeo (1977 : 49-50), Clas (1987 : 348) et López-Rúa (2002 : 44, 46 ; 2004 : 65-66).

<sup>14</sup> Renner cite Davy (2000 : 67), Paillard (2000 : 81), Bauer et Huddleston (2002 : 1635), Bauer (2003 : 47), Gries (2004 : 645-647) et Booij (2005 : 20). Ajoutons Bat-El (2006 : 66).

<sup>15</sup> Il s'agirait de « clipped-compounds » (Bat-El 2006 : 66), soit de composés par troncation.

double troncation. Les amalgames dont les deux bases-sources sont tronquées, comme *alicament* (< aliment + médicament), respectent la contrainte DBLE TRONC.

### 1.2.3. Contrainte sémantique : COORD

La seconde caractéristique définitoire suppose une contrainte sémantique, la coordination (COORD). Renner explique que selon Kubozono (1990 : 1-2), Plag (2003 : 122) et Grésillon (1984 : 120-121) il existe une nuance entre les « blends »<sup>16</sup> de type *motel* (< motor + hotel) et les « proper blends » de type *alicament* (< aliment + médicament) ou *modem* (< modulator + demodulator). Cette contrainte repose sur un lien de coordination entre les signifiés des bases-sources. Ainsi, les « blends » ne répondent pas positivement à l'énoncé-test « un X est à la fois un Y et un Z », tandis que les « proper blends » y répondent positivement. En effet, un *modem* est à la fois un modulateur et un démodulateur, tout comme un *alicament* est à la fois un aliment et un médicament. En revanche, un *motel* n'est pas à la fois un motor-car et un hôtel.

### 1.2.4. Contrainte morpho-phonologique : INTERP

La troisième et dernière caractéristique définitoire mise en évidence par Renner est d'ordre morpho-phonologique : l'interpénétration (INTERP). Selon certains linguistes<sup>17</sup> il est nécessaire qu'un segment de l'amalgame soit commun aux bases-sources. Ce critère exclut donc des mots comme *brunch* (< breakfast + lunch) ou *caméscope* (< caméra + magnétoscope) de la catégorie, contrairement à des mots tels que *motel* pour lesquels il existe un segment commun. Le segment *ot* commun à *motor* et à *hotel* est dit ambimorphémique<sup>18</sup>.

### 1.2.5. Un prototype pour le français?

L'approche proposée par Renner (2006) est celle qui a été adoptée dans le cadre de notre recherche et de cette étude. Il s'agit pour lui de procéder à une approche prototypiste de l'amalgamation lexicale. En réunissant les études concernant le phénomène dans autant de langues que possible, on peut établir un certain nombre de contraintes récursives qui sont à respecter pour considérer l'unité comme un amalgame lexical, et ainsi déterminer un niveau de prototypie. Cette approche constitue un modèle complexe et suffisamment puissant pour déterminer ce niveau de prototypie, mais il faut, à notre sens, affiner ce modèle pour chaque langue. Comme il a été évoqué *supra*, certains facteurs ne concernent que certaines langues et, afin de proposer le meilleur candidat, il faut considérer chaque langue à partir de ce modèle général. C'est ce que nous proposons de faire dans le §3 en utilisant deux corpus complémentaires.

---

<sup>16</sup> Selon la terminologie de Plag (2003).

<sup>17</sup> Grésillon (1984 : 15), Bankov, Dimitrov et Dragnev (1989 : 10) et Piñeros (1999 : 1).

<sup>18</sup> Piñeros (1999 : 14 ; 2004 : 225).

## 2. LES CORPUS

Étant donné qu’au regard des études antérieures sur le phénomène, on peut ne pas être entièrement satisfait des méthodes d’identification et de classification des amalgames lexicaux, nous proposons d’adopter une approche s’appuyant sur un corpus original. En effet, les corpus sur lesquels reposent les études précédentes présentent soit des natures très hétérogènes, soit comportent sur un nombre trop restreint d’items. Le caractère extragrammatical de l’amalgamation lexicale ne doit pas parasiter la cohérence du corpus. Ainsi, Grésillon (1984) et Fradin (1997) s’appuient sur des corpus trop hétérogènes pour tirer des résultats généraux. Galisson (1987) s’appuie sur un corpus intéressant, constitué de mots fantaisistes mais qui ne démontre rien de la langue courante. Clas (1987) utilise des données attestées mais uniquement issues de lexiques scientifiques qui risquent donc de présenter des particularités propres à la terminologie scientifique.

L’utilisation parallèle de deux corpus distincts et complémentaires a été avancée en introduction. Cette section précise la nature de ces corpus. Le premier, que nous avons baptisé COAL<sup>19</sup> est le plus important quantitativement puisqu’il est composé d’environ 30 fois plus d’amalgames que le second. Mais ces corpus ne s’opposent pas uniquement par le nombre de données qu’ils renferment. COAL est constitué d’amalgames lexicaux tirés de dictionnaires fantaisistes, c’est-à-dire de termes forgés *in vitro*, dans le cadre d’un projet lexicographique. Le second corpus est un agglomérat de termes issus de différentes sources – qui vont être détaillées *infra* – et qui ont été forgés *in vivo*, c’est-à-dire dans le cadre de situations discursives.

Le rapport entre ces deux corpus est, dans une certaine mesure, comparable à l’opposition entre grammaticalité et extragrammaticalité. Les amalgames créés *in vivo* répondent à un besoin de communication, tandis que les termes contenus dans COAL sont factices et n’ont pas vocation à apparaître ailleurs que dans les dictionnaires d’où ils sont tirés.

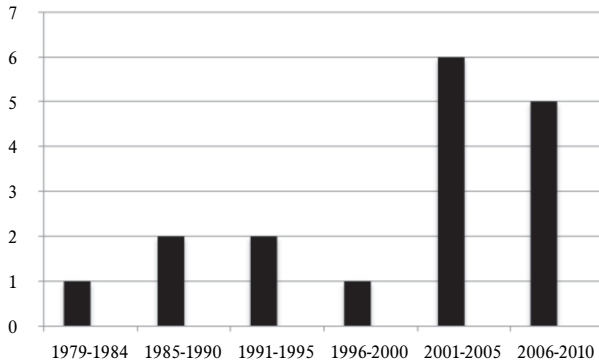
### 2.1. Des dictionnaires fantaisistes aux amalgames lexicaux

COAL est un corpus construit par l’extraction de 7052 amalgames lexicaux de 17 dictionnaires fantaisistes. Les dictionnaires d’amalgames lexicaux utilisés pour constituer COAL appartiennent à une catégorie de dictionnaires particuliers : les dictionnaires fantaisistes<sup>20</sup>. Ce type de dictionnaire est singulier à plusieurs titres. Tout d’abord le nombre de dictionnaires publiés au cours des 30 dernières années est particulièrement important. Depuis 1979, 17 dictionnaires de mots-valises fantaisistes ont paru selon la chronologie illustrée par le graphique 1. Ainsi en 9 ans (2001 à 2010), pas moins de 11 dictionnaires de mots-valises ont vu le jour en France, ce qui est considérable pour ce type de littérature.

---

<sup>19</sup> Pour COrpus d’Amalgames Lexicaux.

<sup>20</sup> Cf. Léturgie (2010).



Graphique 1 : répartition chronologique des dictionnaires d'amalgames depuis 1979

Bien que ces différents dictionnaires répondent principalement à une fonction ludique, ils ne s'y limitent pas tous. Certains rédacteurs de ces dictionnaires nourrissent en effet un objectif pédagogique, et cherchent à utiliser l'amalgamation lexicale comme outil de sensibilisation au lexique du français. C'est notamment le cas de Galisson et Porcher (1986) qui ont conçu le *Distractionnaire* comme un outil didactique à destination des locuteurs étrangers. Nous avons, par ailleurs, eu l'occasion de constater que des ateliers de création de mots-valises sont régulièrement organisés dans les classes de l'enseignement primaire. Il est intéressant de constater que le caractère extragrammatical de l'amalgamation lexicale n'entre pas en conflit avec leur emploi dans un cadre didactique<sup>21</sup>.

COAL est une base de données de 7052 items dont les bases-sources ont été identifiées. Chaque amalgame est catégorisé selon son schéma de construction en suivant les modèles de Clas (1987) et de Grésillon (1984) et étiqueté en fonction des contraintes de Renner (2006).

## 2.2. Un corpus *in vivo*

En regard de COAL, il importe de constituer un corpus de mots forgés *in vivo*. Ce corpus, auquel nous ferons référence par le sigle CALA<sup>22</sup>, comporte aujourd'hui 214 termes issus de différentes sources. La source primaire utilisée pour cette étude est le corpus français de Grésillon (1984 : 176-186) pour 130 items, puis une partie de la base de données *Neologia* pour 56 items et enfin le *Grand Robert* 2009 pour 28 items. CALA n'a pas encore bénéficié d'une analyse sur l'ensemble de ses items et on estime à un peu plus de 250 le nombre d'items qu'il réunira.

<sup>21</sup> Cf. Léturgie (à paraître) pour plus de détails au sujet de l'utilisation didactique d'amalgames lexicaux.

<sup>22</sup> Pour Corpus d'Amalgames Lexicaux Attestés.

Le corpus de Grésillon (1984) a fait l'objet d'un dépouillement afin de ne conserver que les items issus de certaines catégories discursives. Ont donc été conservés les termes marqués comme venant de l'audiovisuel, du commerce, de l'édition, du journalisme, et de la publicité ; les amalgames issus de textes littéraires ou produits par des malades mentaux ont été écartés. Les amalgames bénéficient des mêmes renseignements que dans COAL.

### **3. À LA RECHERCHE D'UN PROTOTYPE POUR LE FRANÇAIS**

Maintenant que les corpus ont été présentés, il est possible de leur appliquer les différentes théories introduites au §1. Les deux corpus seront dans un premier temps interrogés par les approches classiques, c'est-à-dire les quatre approches principales pour le français. Cette première analyse va faire apparaître un schéma d'amalgamation prototypique qu'il faudra corrélérer aux contraintes proposées par Renner (2006).

#### **3.1. Interrogation des corpus par les approches classiques**

Seuls les modèles de Clas (1987) et de Grésillon (1984) vont être appliqués aux corpus. La théorie de Galisson (1987) n'apporte pas de données supplémentaires à ces deux études et va donc être mise de côté. L'analyse menée par Fradin (2000) nécessite une longue recherche et ne peut être présentée dans cet article. Cependant, l'analyse des structures phonologiques des amalgames fantaisistes est un travail que nous souhaitons conduire dans notre thèse.

Les corpus vont livrer des informations sur les structures morphologiques, puis morpho-phonologiques dominantes en français. Nous allons voir que, comme on l'avait postulé, les résultats fournis par COAL coïncident avec ceux de CALA.

##### **3.1.1. Approche morphologique**

La première analyse est d'ordre morphologique. En adaptant la typologie de Clas (1987), on obtient des données sur la morphologie des amalgames lexicaux. Rappelons que Clas (1987) repose sur un corpus de termes techniques, essentiellement de physique ou chimie. Les 6 premiers schémas sont ceux originellement présentés par Clas (1987) tandis que le dernier concerne les cas d'interpénétrations parfaites. Le modèle original ne laissant pas de place à l'aspect phonologique dans la formation des amalgames, il a fallu ajouter ce schéma puisque 1424 items de COAL suivent ce patron.



type d'amalgamation	exemple in vivo	CALA		exemple in vitro	COAL	
		#	%		#	%
(1) apocope + aphérèse	domotique < dom[icile] + [rob]otique	40	18,7%	camembour < cam[embert] + [cal]embour	811	11,5%
(2) apocope + apocope	telex < tel[eprinter] + ex[change]	2	0,9%	∅	0	0,0%
(3) aphérèse +aphérèse	nylon < vinyl + coton	0	0,0%	∅	0	0,0%
(4) apocope simple	publipostage < publi[cité] + postage	87	40,7%	scrypte < s[cript] + crypte	3408	48,3%
(5) aphérèse simple	robotique < robot + [informa]tique	46	21,5%	toasticomane < toast + [tox]icomane	1081	15,3%
(6) apocope/aphérèse + syncope	upérisation < u[ltra] + p[ast]e[u]risation	13	6,1%	escameloter < esca[mot]ter + [ca]melote	328	4,7%
(7) interpénétrations parfaites	∅	26	12,1%	bovidéaliste < bovidé + idéa-liste	1424	20,2%
		<b>214</b>	<b>100%</b>		<b>7052</b>	<b>100%</b>

Tableau 1 : analyse du corpus selon la méthode de Clas (1987)

Le tableau 1 révèle que la majorité des amalgames des deux corpus est formée par apocope simple. Il faut en réalité entendre que c'est la base-source gauche qui est apocopée et que la base-source droite apparaît entièrement. On remarque que les autres schémas sont plutôt proportionnels entre les deux corpus. On remarque cependant que les cas (2) et (3) sont absents de COAL alors que le patron (2) apparaît à deux reprises dans CALA. Si le schéma (3) n'apparaît pas, c'est que le corpus de Clas (1987) n'a pas encore été intégré à CALA. Il est intéressant de constater que COAL ne comporte pas d'unités adoptant la structure (2) ou (3). Cela signifie que la contrainte TRONC INT est respectée par 100% des amalgames fantaisistes. Ce que l'on note également c'est que les 2 termes qui correspondent à la structure (2) proviennent du *GR* et sont des termes techniques empruntés à l'anglais (*algol* < algorithmic language et *fortran* < formular translation). Le schéma ne semble donc pas productif en français dans CALA non plus.

Le problème de cette analyse est que toute considération sur la phonologie des amalgames est écartée. Or, la présence de segments homophones joue un rôle de première importance. Rappelons que Grésillon (1984 : 15) considère que la présence

d'un segment (quasi-)homophone comme « la condition *sine qua non* de la formation du mot-valise ». Fradin, Montermini et Plénat (2009 : 35) signalent qu'« un survol des données disponibles [...] permet de supposer que les mots-valises maximisent en leur sein le nombre et l'étendue des tronçons communs aux lexèmes dont ils sont issus. » C'est pour cette raison qu'il est nécessaire de conduire une analyse d'ordre morpho-phonologique.

### **3.1.2. Approche morpho-phonologique**

En reprenant les schémas d'amalgamation de Grésillon (1984), il est possible de quantifier la proportion d'items créés par troncation ou par enchâssement tout en mesurant la part d'homophonie au sein de chaque matrice. Le segment homophone identifié ici correspond à la contrainte INTERP. Il s'agit donc de ce que Piñeros (1999 et 2004) appelle segment ambimorphémique : un segment homophone et situé au point de jonction des bases-sources. Les segments homophones situés à une autre place dans l'amalgame ne sont pas pris en compte ici. Il serait pourtant intéressant de vérifier la place de ce type de segment à la manière de ce qu'à fait Fradin (2000 : 28-30).

type d'amalgamation	exemple in vivo	CALA		exemple in vitro	COAL	
		#	%		#	%
(8) avec segment homophone	enfantaisie < enfant + fantaisie	26	12,1%	bovidéaliste < bovidé + idéaliste	1427	20,2%
<b>TOTAL homophonie</b>		<b>26</b>	<b>12,1%</b>		<b>1427</b>	<b>20,2%</b>
(9a) avec troncation (aphérèse + apocope)	amarouché < am[ouraché] + [effa]rouché	20	9,3%	urigence < uri[ner] + [ur]gence	191	2,7%
(9b) avec troncation (apocope)	bavardhurler < bavarder + hurler	21	9,8%	coquinlicot < coquin + coquelicot	421	6,0%
(9c) avec troncation (aphérèse)	infiniverti < infini + [in]verti	14	6,5%	ripublicain < ripou + [répu]blicain	335	4,8%
<b>TOTAL troncation</b>		<b>55</b>	<b>25,7%</b>		<b>947</b>	<b>13,4%</b>
(10a) avec segment homophone et troncation (aphérèse + apocope)	falsifis < fal[sifier] + [sal]sifis	20	9,3%	racifisme < ra[ci]sme + [pa]cifisme	621	8,8%
(10b) avec segment homophone et troncation (apocope)	éphémerde < éphé[mère] + merde	68	31,8%	buthym < bu[ti]n + thym	2982	42,3%
(10c) avec segment homophone et troncation (aphérèse)	bisouterie < bisou + [bijou]terie	32	15,0%	rénalcitrant < rénal + [récal]citrant	747	10,6%
<b>TOTAL troncation + homophonie</b>		<b>120</b>	<b>56,1%</b>		<b>4350</b>	<b>61,7%</b>
(11a) avec enchâssement	ubiamourquité < ubiquité + amour	11	5,1%	obchèqueieux < obséqueux + chèque	77	1,1%
(11b) avec enchâssement et homophonie	enseignement < en[seigne]ment + saigne	2	0,9%	désautobéis- sance < désobéis- sance + auto	251	3,6%
<b>TOTAL enchâssement</b>		<b>13</b>	<b>6,1%</b>		<b>328</b>	<b>4,7%</b>
		<b>214</b>	<b>100%</b>		<b>7052</b>	<b>100%</b>

Tableau 2 : analyse du corpus selon la méthode de Grésillon (1984)

Le tableau 2 met plusieurs éléments en évidence. Tout d'abord le rapport entre (8) et (9) est presque inverse entre COAL et CALA. Les schémas de troncation sans homophonie sont préférés au schéma (8) dans CALA alors que les amalgames fantaisistes respectent la contrainte INTERP. Globalement, c'est malgré tout le schéma (10b) qui s'impose dans les deux corpus. Ce schéma correspond simplement au schéma (4) avec respect de la contrainte INTERP, c'est-à-dire que la base-source droite partage au moins un phonème commun avec la base-source gauche.

Les schémas (11a) et (11b) sont, à l'image de ce qu'on vient d'observer entre (8) et (9), en miroir. L'enchâssement sans homophonie est préféré par CALA tandis que COAL compte trois fois plus de termes répondant à (11b).

Les résultats recensés dans le tableau 2 confirment la tendance dégagée par le tableau 1 : la majorité des amalgames lexicaux sont construits par apocope de la base-source gauche et comportent un segment ambimorphémique. On s'aperçoit que cette dernière propriété est importante puisque dans le cas des troncations sans homophonie, le schéma (9c) est à peine plus fréquent que (9a) pour CALA ou que (9c) pour COAL.

### 3.2. Mesure de prototypie des corpus

À partir des données qui ont été réunies par l'analyse des nomenclatures de 17 dictionnaires fantaisistes d'amalgames lexicaux et de leur comparaison avec des données attestées en français, il est possible de déterminer un prototype d'amalgame lexical. À l'aide des tableaux 1 et 2, on peut établir que le prototype répond aux propriétés morpho-phonologiques suivantes :

- l'unité est formée par amalgamation de deux lexèmes ;
- la base-source gauche subie une apocope ;
- la base-source droite reste intacte ;
- un segment homophone est présent au point de jonction des bases-sources.

Ces propriétés correspondent partiellement au schéma (4) et totalement au schéma (10b). Les unités qui ne respectent pas l'ensemble de ces propriétés (comme les termes construits par les différents schémas qui viennent d'être présentés) ne sont pas exclues de la catégorie des amalgames lexicaux mais s'échelonnent sur un degré de typicité moindre.

Si on transpose ces observations avec les contraintes proposées par Renner (2006), on obtient un respect de deux contraintes sur trois puisque nous écartons sciemment la contrainte COORD dans l'analyse proposée ici. COAL montre que dans 61,7% des cas, les contraintes TRONC INT et INTERP sont respectées, et 29%<sup>23</sup> si on ajoute la contrainte DBLE TRONC. En ce qui concerne CALA, 56,1% des amalgames respectent les contraintes TRONC INT et INTERP. Si la contrainte DBLE TRONC est ajoutée, il ne reste que 21,4% du corpus ce qui est inférieur à la proportion d'amalgames

---

<sup>23</sup> On additionne le schéma (8) et le (10a).

ne respectant que TRONC INT<sup>24</sup>. Ce résultat peut surprendre puisqu'on semble s'éloigner du prototype proposé par Renner (2006). En réalité, nous pensons que le français a tendance à favoriser le schéma (10b) et que, même si l'unité résultante semble moins parfaitement construite qu'en suivant le schéma (8), il s'agit bien du prototype pour le français dans la mesure où la majorité des amalgames collectés dans les deux corpus correspondent au schéma (10b).

## CONCLUSION

L'extragrammaticalité morphologique des amalgames lexicaux n'a pas fait l'objet de commentaires autres que les observations de Fradin (2000) et Fradin *et al.* (2009). Nous pensons qu'il est nécessaire de continuer à documenter le phénomène pour développer ces observations, et que cette documentation passe notamment par l'analyse d'un corpus tel que COAL. S'il ne permet pas de transférer ses résultats directement au lexique général du français, l'approche contrastive que nous proposons semble être un bon moyen de confirmer ou d'infirmer certaines propriétés des amalgames lexicaux qui ne constituaient finalement que des tendances sur des ensembles réduits de mots.

Les modèles théoriques présentés au §1 sont d'excellents outils d'analyse. On voit qu'en confrontant les données obtenues par application de ces théories aux deux corpus, un consensus apparaît. Étant donné que les résultats de COAL coïncident avec ceux de CALA, on voit véritablement un prototype, vérifié par un nombre conséquent de données, se dégager. Il n'est pas absurde de considérer le schéma le plus productif comme étant le plus prototypique dans la mesure où on s'appuie sur plus de 7000 items, ce qui représente une base de données suffisante. Nous désirons cependant continuer cette analyse en tenant compte de l'aspect sémantique des amalgames lexicaux. L'approche la plus complète est, avec celle de Grésillon (1984), celle de Fradin (1997, 2000). Il faut donc vérifier si les modèles théoriques de ces chercheurs s'adaptent à nos deux corpus et les résultats que l'on peut en tirer. Une dernière étape dans la mise au jour d'un prototype serait de conduire une étude contrastive sur plusieurs langues, à l'image de ce qu'a fait Renner (2006). Nous postulons en effet que le schéma le plus productif ne sera pas nécessairement le même dans toutes les langues, notamment du fait de la structure accentuelle des unités lexicales ; la phonologie jouant, comme on l'a vu, un rôle crucial dans la formation des amalgames.

Enfin, alors que leur extragrammaticalité morphologique – et notamment leur imprédictibilité – rend les amalgames lexicaux difficiles à détecter automatiquement, la description que nous proposons doit, à terme, permettre de fournir assez d'éléments pour élaborer des outils de détection automatique. Cette détection s'avère complexe du fait de l'extragrammaticalité de l'amalgamation lexicale et ce n'est qu'en accumulant un grand nombre de données qu'il sera possible de développer ce genre d'outils.

---

<sup>24</sup> 25,7% des amalgames de CALA ne respectent que la contrainte TRONC INT alors que le pourcentage tombe à 13,4% dans COAL.

## Références bibliographiques

- ADAMS, Valerie (1973) *An Introduction to Modern English Word-Formation*. Harlow : Longman.
- ALGEO, John (1977) "Blends, a structural and systemic view." *American Speech* 52, 47-64.
- BANKOV, Dimitri/Dimitar DIMITROV/Vladimir DRAGNEV (1989) "Contribution à l'étude des mots-valises ou "mixonymes." *Supostavitelno Ezikoznanie* 14 (5), 8-11.
- BAT-EL, Outi (1996) "Selecting the best of the worst: the grammar of Hebrew blends." *Phonology* 13 (3), 283-328.
- BAT-EL, Outi (2006) "Blends." In K. Brown (éd.), 66-70.
- BAUER, Laurie (2003) *Introducing Linguistic Morphology*, 2<sup>nd</sup> ed., Edinburgh : Edinburgh University Press.
- BAUER, Laurie/Rodney HUDDLESTON (2002) "Lexical Word-formation." In : R. Huddleston/G. Pullum (éds), 1621-1721.
- BERTINETTO, Pier Marco (2001) "Blends and syllable structure: A four-fold comparison." In : Mercé Lorente et al. (éds), 59-112.
- BOOIJ, Geert (2005) *The Grammar of Words*. Oxford : Oxford University Press.
- BOOIJ, Geert/Jaap VAN MARLE (éds) (1990) *Yearbook of Morphology* vol. 3, Dordrecht : Foris Publications.
- BRDAR-SZABÓ, Rita/Mario BRDAR (2008) "On the marginality of lexical blending." *Jezikoslocije* 9/1-2, 171-194.
- BROWN, Keith (éd.) *Encyclopedia of Language and Linguistics*, 2<sup>nd</sup> ed., Vol. 2. Oxford : Elsevier.
- CLAS, André (1987) "Une matrice terminologique universelle : la brachygraphie gigogne." *Méta* 32 (3), 347-355.
- CORBIN, Danielle et al. (éds) (1997), *Sillexicales 1, Mots possibles et mots existants*. Villeneuve d'Ascq : URA 382 - CNRS/Université de Lille 3.
- DAVY, Dennis (2000) "Shortening Phenomena in Modern English Word Formation: An Analysis of Clipping and Blending." *Franco-British Studies* 29, 59-76.
- DOLESCHAL, Urusula/Anna M. THORNTON (éds) (2001) *Extragrammatical and Marginal Morphology*. München : Lincom Europa.
- DYKSTRA, Anne/Tanneke SCHOONHEIM (éds) (2010) *Proceedings of the XIV Euralex International Congress*. Ljouwert : Afûk/Fryske Academy.
- FRADIN, Bernard (1997) "Les mots-valises : une forme productive d'existants impossibles ?" In : D. Corbin et al. (éds), 101-110.
- FRADIN, Bernard (2000) "Combining forms, blends and related phenomena." In U. Doleschal /A. M. Thornton (éds), 11-59.
- FRADIN, Bernard/Fabio MONTERMINI/Marc PLÉNAT (2009) "Morphologie grammaticale et extragrammaticale." In : B. Fradin et al. (éds), p. 21-45.
- FRADIN, Bernard/Françoise KERLEROUX/Marc PLÉNAT (2009) *Aperçus de morphologie du français*. Saint-Denis : Presses Universitaires de Vincennes.
- GALISSON, Robert (1987) "Les dictionnaires de parodie comme moyens de perfectionnement en langue française." *Études de linguistique appliquée* 67, 57-118.
- GRÉSILLON, Almuth (1984) *La Règle et le monstre : le mot-valise. Interrogations sur la langue, à partir d'un corpus de Heinrich Heine*. Tübingen : Niemeyer.
- GRIES, Stefan Th. (2004) "Shouldn't it be breakfunch? A quantitative analysis of blend structure in English." *Linguistics* 42 (3), 639-667.

- HUDDLESTON, Rodney/Geoffrey PULLUM (éds) (2002) *The Cambridge Grammar of the English Language*. Cambridge : Cambridge University Press.
- KUBOZONO, Haruo (1990) "Phonological Constraints on Blending in English as a Case for Phonology-Morphology Interface." In G. Booij/J. Van Marle (éds), 1-20.
- LÉTURGIE, Arnaud (2010) "Une pratique lexicographique émergente : les dictionnaires détournés." In : A. Dykstra/T. Schoonheim (éds), 1340-1346.
- LÉTURGIE, Arnaud (2011) "À propos de l'amalgamation lexicale en français." *Langages* 183, 75-88.
- LÉTURGIE, Arnaud (à paraître) "Des dictionnaires fantaisistes d'amalgames lexicaux comme dictionnaires d'apprentissage." *Nouvelles perspectives en lexicographie*.
- LÓPEZ RÚA, Paula (2002) "On the Structure of Acronyms and Neighbouring Categories: A Prototype-Based Account." *English Language and Linguistics* 6 (1), 31-60.
- LÓPEZ RÚA, Paula (2004) "The Categorical Continuum of English Blends." *English Studies* 85 (1), 63-76.
- MERCÉ, Lorente et al. (éds) (2001) *La gramàtica i la semàntica en l'estudi de la variació*. Barcelona : Promociones y Publicaciones Universitarias.
- PAILLARD, Michel (2000) *Lexicologie contrastive anglais-français. Formation des mots et construction du sens*. Gap : Ophrys.
- PEREIRO, Myriam /Henry DANIELS (éds) (2006) *Le Désaccord*. Nancy : Publications de l'AMAES.
- PIÑEROS, Carlos-Eduardo (1999) "Word-Blending as a Case of Non-Concatenative Morphology in Spanish." *Rutgers Optimality Archive* 343-0999, <<http://roa.rutgers.edu>>, consulté le 13 septembre 2010.
- PIÑEROS, Carlos-Eduardo (2004) "The Creation of Portmanteaus in the Extragrammatical Morphology of Spanish." *Probus* 16 (2), 203-240.
- PLAG, Ingo (2003) *Word-Formation in English*. Cambridge : Cambridge University Press.
- RENNER, Vincent (2006) "Dépasser les désaccords : pour une approche prototypiste du concept d'amalgame lexical." In : M. Pereiro/H. Daniels (dirs), 137-147.
- SABLAYROLLES, Jean-François (1997) "Néologismes : une typologie des typologies." *Cahiers du C.I.E.L. 1996-1997*, 11-48.
- SABLAYROLLES, Jean-François (2000) *La néologie en français contemporain. Examen du concept et analyse de productions néologiques récentes*. Paris : Champion.
- TOURNIER, Jean (1985) *Introduction descriptive à la lexicogénétique de l'anglais contemporain*. Paris : Champion.

### **Dictionnaires d'amalgames lexicaux**

- CHIFLET, Jean-Loup (1999) *Le Cafard laqué. Les mots-portemanteaux*. Paris : Mots & Cie.
- CHIFLET, Jean-Loup (éd.) (2002) *Mais que fait l'Académie ? Le dictionnaire des mots qui devraient exister*. Paris : Mots & Cie.
- CHIFLET, Jean-Loup/Nathalie KRISTY (1992) *Le Dictionnaire des mots qui n'existent pas*. Paris : Presses de la cité.
- CHIFLET, Jean-Loup/Nathalie KRISTY (2003) *Le Mokimanké : le dico des mots qui existent enfin !* Paris : Mots & Cie.
- CLÉMENT, Yves-Marie/Gérard GRÉVERAND (1993) *Pianissimots, petit dictionnaire des mots-valises*. Levallois Perret : Many.

- CRÉHANGE, Alain (2004) *Le Pornithorynque est un salopare. Dictionnaire de mots-valises*. Paris : Mille et une nuits.
- CRÉHANGE, Alain (2006) *L'Anarchiviste et le Biblioteckel. Dictionnaire de mots-valises*. Paris : Mille et une nuits.
- FINKIELKRAUT, Alain (2006 [1979]) *Petit dictionnaire illustré. Les mots qui manquent au dico*. Paris : Seuil.
- FOURNEL, Paul/Henri CUECO (2007) *Les Animaux d'amour et autres sardinosaures*, Bordeaux : Le Castor Astral.
- GALISSON, Robert/Louis PORCHER (1986) *Distractionnaire*. Paris : Clé international.
- GUILLAUME, Jean-Claude (2008) *La valise diplodocumatique*. Saint-Denis : Édilivre-Aparis.
- LE TELLIER, Hervé/Xavier GORCE (2007) *Les Opossums célèbres*. Bordeaux, Le Castor Astral.
- LOUBIÈRE, Sophie/Olivier LATYK (2003) *Éléphanfare. Le jeu des mots-valises pour les petits*. Paris : Albin Michel Jeunesse.
- MINDA, Alexandre (2004) *Mots-valises à poignées*. Paris : L'Harmattan.
- ROUBAUD, Jacques (2008) *Sardinosaures & Cie*. Bourges : Mille univers.
- SERGUINE, Jacques (1988) *Contradictionnaire*. Paris : Ramsay.
- THIBAUD, Jean-Jacques (2005) *Nouveaucabulaire*. Paris : Recherche midi.

#### Résumé

#### UN CAS D'EXTRAGRAMMATICALITÉ PARTICULIER : LES AMALGAMES LEXICAUX FANTAISISTES

L'amalgamation lexicale est un procédé d'innovation lexicale assez peu décrit en français et dont la morphologie ressort de l'extragrammatical. Partant du constat que les différentes classifications des amalgames lexicaux sont établies sur des corpus restreints, nous proposons de dresser un état des lieux de ces classifications en confrontant plusieurs modèles théoriques avec deux corpus complémentaires pour fournir une description morpho-phonologique de ces unités lexicales afin de faire émerger un prototype pour le français. Nous avons procédé, dans un premier temps, au recensement de 7 052 amalgames lexicaux tirés de 17 dictionnaires fantaisistes. Puis, un second corpus d'amalgames attestés en discours a été constitué afin d'être comparé au premier. À l'aide des données récoltées, il est possible de voir apparaître un schéma prototypique de l'amalgamation lexicale en français.

#### Povzetek

#### POSEBEN PRIMER IZVENSISTEMSKOSTI : IZMIŠLJENI LEKSIKALNI AMALGAMI

Leksikalno amalgamiranje je postopek izmišljanja besed, ki je v francoščini dokaj slabo raziskan in ki sodi na področje izvensistemske morfologije. Ugotavljamo, da so se različne tipologije leksikalnih amalgamov do sedaj opirale na korpuse manjšega obsega. V članku tako podamo pregled dosedanjih tipologij, več teoretskih modelov pa soočimo z gradivom dveh korpusov, s čimer pridemo do morfofonološkega opisa takšnih leksikalnih enot, kar nam omogoča izluščiti prototip francoskih leksikalnih amalgamov. Najprej smo sestavili korpus 7052 leksikalnih amalgamov, ki smo jih našli v 17 slovarjih izmišljenih besed. Nato smo zgradili korpus amalgamov, ki smo jih našli v dejanski rabi in ga primerjali s prvim. Na podlagi pridobljenih podatkov je moč določiti prototipsko shemo leksikalnega amalgamiranja v francoščini.



## ON UNFAMILIAR ITALIAN LEXICAL BLENDS FROM NAMES AND NOUNS

### 1. INTRODUCTION

Italian has recently witnessed a steady increase in the use of unfamiliar lexical blends from names and nouns with an identificatory and descriptive function (in the sense of Anderson 2007), e.g. *Draquila* < *Dracula* [Dracul] + *L'Aquila*, or *docciabatta* < *doccia* [shower] + *ciabatta* [slipper]. An important fact about blends is that they are created in extragrammatical morphology (Dressler 2000; Ronneberger-Sibold 2006, 2010). Unlike regular inflection and word formations – governed by rule-based grammatical competence –, they are consciously created to reduce morphotactic and morphosemantic transparency (Dressler et al. 1987; Dressler 1999). The resulting increase in processing time and effort makes the blend memorable (Lehrer 2003) in the interest of relevance (Sperber/Wilson 1986/1995) and effectiveness (maximum success chances) versus efficiency (minimum effort) (cf. de Beaugrande/Dressler 1981). More particularly, while intentional blends from names and nouns are coined with careful attention to the semantic concepts encoded by the individual source words (SWs), processing and understanding the blend depends on the decoder's direct or surrogate experience of the related reference.

Whereas lexical blending has typically been dismissed as an elusive and peripheral word-formation process cross-linguistically, it is found to be underrepresented in Italian (as compared to English, German and French, cf. Bertinetto 2001). It is therefore not surprising that blending has been often neglected in Italian linguistics. As far as we are aware, the only exceptions are Migliorini (1949), Thornton (1993, 2004b) and Bertinetto (2001). And yet, it has received far less attention than it deserves. This paper is thus an attempt to redress the imbalance and contribute to the (by now growing) body of work on morphological blending. Our main concern relates most directly to the morphosyntactic and morphosemantic transparency – and, ultimately, conceptual motivation – of Italian blends from names and nouns as operations of extragrammatical morphology.

The paper is structured as follows. Because the delimitation of the category is a matter of debate, in Section 2 we provide a working definition of lexical blends. At the same time, we touch upon the identificatory function of names. Section 3 is devoted to our research question. Specifically, we shall carry out a qualitative investigation into a restricted number of examples selected from a collection of 100 blends that we mainly gathered from the World Wide Web. They are used in media language and journalese, radio programmes, TV

---

\* *Author's address:* Università degli Studi di Modena e Reggio Emilia, Dipartimento di Scienze del Linguaggio e della Cultura, Largo Sant'Eufemia 19, 41121 Modena, Italia. Email: [silvia.cacchiani@unimore.it](mailto:silvia.cacchiani@unimore.it)

programmes and movies, product advertising and marketing, and event advertising. While Section 3.1 concentrates on morphotactic transparency, Section 3.2 deals with morphosemantic transparency and conceptual motivation. Naturalness Theory (Dressler et al. 1987, Dressler 1999) is extremely well-designed to analyze and categorize blends, with an eye at their categorical status within the radial category of subtractive word formations. We shall further touch upon conceptual motivation in the composite structure of individual morphological blends using concepts from Conceptual Metaphor Theory (CMT, cf. Lakoff/Johnson 1980) and Blending Theory (Ruiz de Mendoza 1998). This will enable us to provide some preliminary remarks on the meaning predictability of specific subtypes, and blends from personal names used in journalese and media language in particular (e.g. *Berlusconi* < *Berlusconi* + *Bertinotti*, *Berlingotti* < *Berlinguer* + *Bertinotti*), from the perspective of Cognitive Linguistics (cf. Langacker 1987, Kemmer 2003).

## 2. BLENDS FROM NAMES AND NOUNS

On a narrow definition of the term (Bat-El 2006: 66), lexical blends are word formations in which the inner edges of the SWs are truncated, which excludes clipped compounds (*sitcom* < *situation* + *comedy*, *mocamp* < *motor* + *camp*). The most widely accepted view, summarized by Gries (in press), is that an intentional lexical blend is:

an intentional fusion of typically two (but potentially more) words where a part of a first source word ( $sw_1$ ) – usually this part includes the beginning of  $sw_1$  [left base word] – is combined with a part of a second source word ( $sw_2$  [right base word]) – usually this part includes the end of  $sw_2$  – where at least one source word is shortened and/or the fusion may involve overlap of  $sw_1$  and  $sw_2$ ,

which excludes speech error blends and clipped compounds (Gries, in press: *complex clippings*).

Following Migliorini (1949), Italian blends are typically addressed under the heading *parole macedonia* (literally, Eng. fruit-salad words, cf. Thornton 1993: 153).<sup>1</sup> Italian *parole macedonia* (Thornton 2004b) appear to be mainly coined juxtaposing SW beginning and SW, or beginning plus beginning. A few recent creations are:

- splinter<sup>2</sup> + SW: *docuromanzo* < *documentario* [documentary]/*docu*<sup>3</sup> + *romanzo* [novel] (possibly based on *docufiction*, a borrowing from Eng. *docu-fiction* < *documentary* + *TV fiction*, cf. GRADIT: DOCUFICTION);

<sup>1</sup> More recent terms for *blend* are *incrocio* [intersection] and *fusione* [blending] (Bertinetto 2001: 62, footnote 1).

<sup>2</sup> With Lehrer (2003), we define splinters as parts of truncated words that do not necessarily coincide with syllables or morphemes.

<sup>3</sup> We specify ‘combining form’ for combining forms that started out in a blend and are now recorded in the *Grande Dizionario Italiano dell’Uso* (GRADIT) under separate entries, and/or are listed among Italian ‘elementi formativi’ [combining forms] in Grossman/Rainer (eds) (2004). Where the form is not recorded in the dictionary and yet it already appears to be forming new word families by analogical extension from the blend, the combining form follows the corresponding SW.

- splinter + splinter: *Edilplast* < *edilizia* [building trade]/combining form *-edil-* + *plastica* [plastic]; *Enervit* < *energia* [energy] + *vitamina* [vitamin]/combining form *-vit*; *Cerealvit* < *cereali* [cereals] + *vitamina* [vitamin]/combining form *-vit*; *colf* < *collaboratrice* [assistant] + *familiare* [family, A]; *Tipolito* < *tipografia* [printer's]/combining form *-lito-* + *litografia* [litographer's]/combining form *-tipo-*).
- SW + splinter: *Mentadent* < *menta* [mint] + *denti* [teeth].

By contrast, word-internal shortening and partial overlap would be features of English blends (e.g. *brunch* < *breakfast* + *lunch*, *motel* < *motor* + *hotel*, or *smog* < *smoke* + *fog*). While cross-linguistic comparison shows that these appear to be highly dispreferred techniques in Italian (Bertinetto 2001: 66), they are nevertheless used in a number of novel Italian blends as adjectives and, most often, nouns and names, e.g.:

- with overlap: *semplogica* < *semplice* [simple/plain] + *logica* [logical]/combining form *-logica*; *gastronauta* < *gastronomo* [gourmet]/neoclassical compound *gastro-* + *astronauta* [astronaut]/combining form *-nauta*, *oroschifo* < *oroscopo* [horoscope] + *schifo* [disgust]; *Sarlusconi* < *Sarkozy* + *Berlusconi*;
- without overlap: *modtro* < *moderno* [up-to-date] + *rétro* [retro]; *radicalsauro* < *radicale* [politically radical] + *dinosauro* [dinosaur]/combining form *-sauro*.

Bauer (1983: 233) loosely defines blends as “new lexeme[s] formed from parts of two (or possibly more) other words in such a way that there is no transparent analysis into morphs”, though “in many cases some kind of analysis can be made” and “at least one element is transparently recoverable”. Blends thus comprise so-called portmanteau words, which exhibit truncation at inner boundaries and optional overlap at the switchpoint (*chunnel* < *channel* + *tunnel*, *guestimate* < *guess* + *estimate*, *mimsy* < *miserable* + *flimsy*). Less core items tend to shade off and merge into word formations such as clipped compounds or extended acronyms when: SW<sub>1</sub> is clipped and SW<sub>2</sub> remains unaltered (*cremains* < *cremate* + *remains*, *mocamp* < *motor* + *camp*); the blend might be processed in terms of a neoclassical compound (*arcology* < *architectural* + *ecology*) or a combining form that started out in blends is present (e.g. *-burger*, *-rama*, *-teria*); or both SWs are truncated at the right edge (*linac* < *LINear* + *ACcellerator*) (all examples from Bauer 1983: 233-238).

In a similar vein, Thornton (1993) distinguishes between prototypical blends, with truncation and overlap, and various types of partial blends – or, better, blends that merge into other categories – which are created in extragrammatical morphology via other techniques, whereby: the final open syllable of SW<sub>1</sub> is deleted via haplology in case of overlap with the onset of the initial syllable of SW<sub>2</sub> (*eroicomico* < *eroico* [heroic] + *comico* [funny]); SW<sub>2</sub> remains unaltered while the truncated word has achieved morphological status as a quasi-prefix (*palaghiaccio* < combining form *pala-/palazzo* [palace/building] + *ghiaccio* [ice]); both SWs are truncated at the right

edge and the two splinters exceed one-syllable length (*Interpol* < *internazionale* [international] + *police* [polizia]/combining form *-pol*).

Whereas on a strict definition of the term blends and neighbouring categories appear to represent a cluster of related phenomena which exhibit family resemblance (Brdar-Szabó/Brdar 2008) within the radial category of subtractive word formations (López-Rúa 2002: *shortenings*), on a broad definition the related phenomena we mentioned above can be seen as less core blends, which differ from more core items along specific scales. Accordingly, Ronneberger-Sibold (2006) regards blending as deliberate extragrammatical compounding and scales German blends and clipped compounds (in her terms, *fragment blends*) by their morphotactic transparency.

Dealing as we do with Italian blends from names and nouns, we adopt Gries's (in press) definition of lexical blends for relatively more core items and allow for relatively less core items, also shading off into the neighbouring category of clipped compounds. The main emphasis, however, lies in the morphotactic transparency of relatively more core items, and, second, in morphosemantic transparency and conceptual motivation.

## 2.1. Identification and description

The approach outlined in this paper proceeds from the observation that blends from nouns and names are variously created for identifying and descriptive purposes. Along with prototypical personal names, bynames, place names, product names, names of abstract notions and names of numbers, days, months, etc., nicknames belong to the category *Name* (vs. *Noun*). Unlike nouns, they serve a *referential identificatory* function in referring to individuals when used as arguments in predications. Unlike nouns or pronouns, names have an indexical in their associated concept, thus exemplifying *direct reference*. Although names may undergo de-semantization and lose memory of their original descriptive, lexical meanings over time, they retain their original identificatory function (Anderson 2007).

Blends from names (and nouns) as names may (but do not necessarily) have a descriptive meaning to them as a result of the etymology of their source names or, more simply, may be created or chosen for their seemingly nice sound effect (e.g. *Josalind* < *Jocelyn* + *Rosalind*, *Tyrus* < *Tyrone* + *Cyrus*, cf. Marchand 1969). In other cases (e.g. celebrity (nick-)names such as *Brangelina* < *Brad (Pitt)* + *Angelina (Jolie)*), however, the choice is determined by *description* and blends are coined with careful attention to the semantic concepts related to the individual SWs. Understanding the blend does not rest exclusively on the recognizability of the two SWs, but correlates with their discriminatory, classificatory and expressive/evaluative functions, which are ultimately responsible for their denotations and connotations. *Description* has a role in fixing the referent, though not in making reference (cf. Kripke 1972/1981). Knowledge of the referent of a name is through direct or surrogate experience of the reference, by ostension or description in its *baptism* (Kripke 1972/1981; Lyons 1977: 217-218: *nomination*).

### 3. DATA ANALYSIS

Blending is not governed by rule-based grammar. Though the SWs are not fully nor immediately accessible to the decoder, the phonological make-up of the blend (size, syllable structure, segmental make-up, etc.) is not accidental: while SW<sub>1</sub> and SW<sub>2</sub> are still *similar* in various ways, similarity of blend and SWs (Gries, in press: *recognizability*) facilitates recognition. (For more on this point, see Bat-El 2006 and references therein; Gries, in press and previous work.) In Section 3.1, we'll be dealing with the recognizability of novel blends from nouns and names within the framework of Natural Morphology (Dressler et al. 1987; Dressler 1999). The *markedness* (as against *naturalness*) of lexemes illustrating different blending techniques will be assessed along the gradable semiotic parameter of *morphotactic transparency* (or morphological constituency of the morphs). *Morphosemantic transparency* (semantic compositionality of the constituent parts) (Dressler et al. 1987; Dressler 1999) and cognitive motivation will be touched upon in Section 3.2, where we provide some preliminary remarks on the role of knowledge in understanding this type of blends.

#### 3.1. Morphotactic transparency

Competing factors which facilitate SW recognition in various ways comprise, among others, length and currency of the truncated words, as well as (Gries, in press, and references therein) a tendency of SW<sub>1</sub> to be more frequent than SW<sub>2</sub>. Blends undergo reduction after being generated as partially predictable novel compounds. Whereas coiners tend to shorten words at their psycholinguistically motivated uniqueness points (with truncation of SW<sub>1</sub> nearly exactly at the selection point and truncation of SW<sub>2</sub> half a phoneme too early in SW<sub>2</sub>, cf. Gries, in press) so as to facilitate analyzability, the semantic plausibility of the connective relation has a bearing on the morphosemantic transparency of the blend and favours identification of the non-head word. Thus, iconic blends based on a coordinative compound (*tigone* < *tigre* [tiger] + *leone* [lion], describing a hybrid (Thornton 1993), or, in our sample, *Berlingotti* < *Berlinguer* + *Andreotti*, describing a political coalition) are less marked than blends that instantiate other relations (*Eataly* < *eat* + *Italy*). Last, the presence of graphemic, phonological and morphemic overlap also works towards analyzability.

Perceptual salience has a major role in determining morphotactic transparency. Broadly speaking, *perceptual salience* (Dressler 1987: 117) determines a preference for syllable-initial as against syllable-final consonants, as well as for suffixes over prefixes (which move roots/stems/lexical bases away from the beginning of the word). Additionally, word beginnings and word ends are more perceptually salient than word-internal segments, which excludes blends such as *\*torhot* (as against *motel* < *motor* + *hotel*). If this is so, then:

- i. In blends whose size (number of syllables) is identical to one of the SWs: we expect SW<sub>2</sub> to be longer (in terms of syllables) than SW<sub>1</sub>, which is more perceptually salient and thus more easily analyzable (*Prodinotti* < *Prodi* + *Bertinotti*, *Happyfania* < Eng. *happy* + *Epifania* [Twelfth Night]).

Among the factors which operate to move the shorter SW to the right are grammatical category of the base words and/or graphemic and phonemic similarity of segments within the blend. E.g.: *Meditathe* < *medita/meditate* [reflect, 2<sup>nd</sup> sing./pl. Imperative] + *thè* [tea, variant from Fr. *thé* of It. *tè*], which retains the stress pattern of SW<sub>1</sub> (with stress on the antepenultimate syllable) and is based on a VN compound with instrument N.

In polysyllabic blends which are identical to the size of both SWs: we expect SW<sub>1</sub> not to be longer than SW<sub>2</sub> (*morbistenza* /morbi'stentsa/, where the onset of the stressed syllable is taken from SW<sub>2</sub>, vs. \**morbidenza* < *morbidezza* /morbi'dettsa/ [softness] + *resistenza* /resi'stentsa/ [strength]).

- ii. In polysyllabic overlap blends: we expect alignment at word edges (in Optimality Theory (OT), cf. Piñeros 2004: 215: ALIGN-MWd) and syllable boundaries. E.g., *Happyfania* /ɛppi'fanja/ results from superimposition of the most likely pronunciation for Eng. *happy* /'hæpi/, It. /'ɛppi/ (GRADIT: HAPPY)<sup>4</sup> – with silent segment *h*, open-mid front vowel, and geminates – and *epifania* /epi'fanja/.
- iii. To facilitate recognition of the less perceptually salient SW<sub>2</sub>, we predict the similarity-motivated preservation of SW<sub>2</sub>'s recognizability to dominate recognizability-motivated preservation of letters and phonemes. Specifically, in blends with word-internal truncation:
- a. If blend's size is identical to both SWs, the so-called size constraint (in OT, cf. Bat-El/Cohen, in press: FAITH METRICAL STRUCTURE (FAITHMS)), would predict that metrical structure (number of syllables and stress pattern) of the blend and SWs are identical (*Berluscotti* < *Berlusconi* + *Bertinotti*; *mordistenza* < *morbidezza* [softness] + *resistenza* [strength], both with stress on the penultimate syllable). Under the principle of saliency, however, the stressed syllable and stress pattern in the blend correspond to the stressed syllable and stress pattern of SW<sub>2</sub> (Bat-El/Cohen, in press: position-based stress assignment. More technically: FAITHHEAD<sub>R</sub> >> FAITHHEAD<sub>L</sub>): *gastro-nauta* /gas'tronawta/ < *gastronomo* /gas'trɔnomo/ [gourmet]/neoclassical compound *gastro-* + *astronauta* /astro'nawta/ [astronaut]/combining form *-nauta* /'nawta/.<sup>5</sup>
- b. If blend's size is identical to the size of one SW, the stress pattern of that SW has priority over the other SW: *britaliano* /brita'ljano/ < *British* / briti \*/ (GRADIT: BRITISH) + *italiano* /ita'ljano/ [Italian]; *oroschifo* /o'rɔskifo/ < *oroscopo* /o'rɔscopo/ [horoscope] + *schifo* / skifo/ [disgust].

<sup>4</sup> All phonetic transcriptions of base nouns, English borrowings and names recorded in the dictionary, are taken from the *Grande Dizionario Italiano dell'Uso* (GRADIT).

<sup>5</sup> Bat-El/Cohen (in press) show that there is no priority between size- and position-based stress assignment in English blends whose size is identical to the size of the left base word. Although our data do not suggest intra-word variation here, our sample is too small to exclude non-crucial ranking between FAITHMS (size) and FAITHHEAD<sub>R</sub> (position) in blends with longer SW<sub>1</sub>.

- c. When blend's size differs from both SWs, if SW<sub>2</sub> is not truncated, SW<sub>2</sub> dominates (Bat-El/Cohen, in press: FAITHHEADW<sub>R</sub> >> FAITHHEADW<sub>L</sub>), as in *radicalsawro* /radical'sawro/ < *radical* /radi'cale/ [radical] + *dinosawro* /dino'sawro/ [dinosaur]/combining form -*sawro* /'sawro/.

### 3.1.1. A typology of Italian blends

Under the *principle of saliency*, Italian bars maximally opaque end-end concatenations. Other blending techniques are available to Italian. Framing our typology within Naturalness Theory and adapting Ronneberger-Sibold (2006) we distinguish three blending techniques, which range from relatively more morphotactically transparent to relatively opaque: i. *complete blending* (*telescope blending* and *inclusive blending*), or the most easily analyzable into constituents; ii. *contour blending*, where blend size is identical to one of the SWs; iii. *semi-complete blending*, with one truncated word and one full SW. Ronneberger-Sibold (2006) also devises another category, *fragment blending* (Gries, in press: *complex clipping*), which combines SW beginnings in clipped compounds. Based on insights from research in Optimality Theory (Bat-El/Cohen, in press) and on their lack of transparency, we suggest viewing *fragment blending* as a neighbouring, though separate category along the continuum from relatively transparent to opaque (cf. Section 3.1.2).

Broadly speaking, although overlap of graphemes and/or phonemes across words represents a marked option (while *biuniqueness*, or one-form one-meaning correspondence, is its opposite on the naturalness scale, cf. Dressler et al. 1987), the more the overlap, the more transparent the blend.

#### Type ia. TELESCOPES

Within *complete blending*, *telescope (syntagmatic) blends* (Algeo 1977) are created via overlap or truncation juxtaposing at converging edges SWs that are initially generated as subsequent words within a phrase. Telescopes are highly infrequent in the corpus, possibly because they represent the most transparent (and, thus, the less memorable and effective) type. Transparency correlates positively with overlap of graphic and sound shapes and, most importantly, morphemes, e.g.: *RavennAntica* /ra'vennan'tika/ < *Ravenna* /ra'venna/ + *antica* /an'tika/ [historical]; *Cooptima* /kɔ'optima/ < *coop* /kɔ'op/ [co-op, originally from It. *cooperativa*] + *ottima* /'ɔttima/ (also It. *optimum* + -*a* fem. sing., or Lat. *optimus*, -*a*, -*um*); or the more peripheral member *Futurauto* < *futura* [next generation, A] + *auto* [car], based on the NP *auto futura/del futuro* [next generation car].

#### Type ib. INCLUSIVE BLENDS

*Inclusive blends* are a feature of written language. They are *associative* (or *portman-teau* cf. Algeo 1977) *overlap blends* in which one constituent includes the other phonologically though not graphemically (Ronneberger-Sibold 2006: 167). Examples from our collection only comprise blends with foreign monosyllabic words, mainly based on exocentric VN compounds, e.g.: *EATaly* < *eat* + *Italy* for an Italian slow food

market, restaurant and now restaurant brand; or the less transparent *Meditathe* /medi'tate/, trade mark for Italian literary coffee mugs, where the blend retains stress pattern and stressed syllable of SW<sub>1</sub> to facilitate analyzability (in line with point iib, Section 3.1). Another example is *Happyfania*, a reduction from AN compound, where phonological similarity replaces phonological identity in the pretonic segment (/ɛppi'fanja/ vs. /epi'fanja/ (cf. Section 3.1, point ii).

#### Type ii. CONTOUR BLENDS

This type comprises blends which retain stress pattern and stressed syllable of one SW (*matrix word*, cf. Ronneberger-Sibold 2006), which is not entirely present in the blend, while the non-matrix can be inserted in full. Generally speaking, adopting the stress pattern of the matrix word (most often SW<sub>2</sub>) facilitates recognition. Other factors that contribute transparency are identical size of blend and matrix word, maximization of SW<sub>1</sub>-SW<sub>2</sub> overlap, inclusion of the SW in the pretonic part of the blend or in the rhyme. Since overlap at word end is relatively marked, inclusion of material from SW<sub>2</sub> in the onset of the stressed syllable might facilitate recognition.<sup>6</sup> Some examples:

- Blend's size and stress pattern are identical to the size and stress pattern of both SWs; size-based stress assignment: *Draquila* /'drakwila/ < *Dracula* /'drakwla/ [Dracul] + (*L*)*Aquila* /'lakwila/; *morbistenza* /morbi'stentsa/ < *morbidezza* /morbi'dettsa/ [softness] + *resistenza* /resi'stentsa/ [strength]; *Berluscotti* /berlu'skotti/ < *Berlusconi* /berlu'skoni/ + *Bertinotti* /berti'notti/; *Itagnolo* /itan'ɲolo/ < *Italiano* /ital'jano/ [Italian] + *Spagnolo* /span'ɲolo/ [spanish].
- Blend's size is identical to the size of both SWs; position-based stress assignment: *gastronauta* /gas'tronawta/ < *gastronomo* /gas'trɔnomo/ [gourmet]/neoclassical compound *gastro-* + *astronauta* /astro'nawta/ [astronaut]/combining form *-nauta* /'nawta/ (cf. Section 3.1, point iia).
- Left-hand alignment of SWs in blends whose size is identical to the size of one SW; size-based stress assignment: *mocaccino* /mokat'tʃino/ < *moca* /moka/ + *cappuccino* /kapput'tʃino/ (without overlap); *frappuccino* /frapput'tʃino/ < *frappé* /frap'pe\*/ + *capuccino* /kapput'tʃino/ (with overlap).<sup>7</sup>
- Right-hand alignment of SWs in blends with overlap whose size is identical to the size of one SW; size-based stress assignment: *Enterogelmini* < *Enterogelmina* /,ɛnterodʒel'mi-

<sup>6</sup> Monosyllabic blends in Italian are loanwords from other languages, e.g. Eng. *smog* or *brunch*, which are not (or are not any longer) analyzed as complex words that combine the inner part of two SWs. Blend size is identical to the size of one SW in complete blends and of both SWs or one of the SWs in contour blends. When blend size differs from both words, the blend retains the stressed syllable of the matrix word, e.g. *Biografilm* /biogra'film/ < *biografia* [biography] /biogra'fia/ + *film* /'film/.

<sup>7</sup> Because SW<sub>1</sub> is monosyllabic and aligns with the stressed syllable of SW<sub>2</sub>, the distinction between size-based and stress-based assignment is not relevant in *Billary* /'billari/ < *Bill* (*Clinton*) /'bil/ + *Hillary* (*Rodham Clinton*) /'illari/



ni/ < *Enterogermina* /ɛnterodʒer'mina/ + *Gelmini* /dʒel'mini/; *oroschifo* /o'rɔskifo/ < *oroscopo* /o'rɔskopo/ [horoscope] + *schifo* /'skifo/ [disgust].<sup>8</sup>

- Blends without overlap whose size is identical to the size of one SW; position-based stress assignment: *Berlingotti* /berlin'gotti/ < *Berlinguer* /berlin'gwer/ + *Andreotti* /andre'otti/.

### Type 3. SEMI-COMPLETE BLENDS

In semi-complete blends the truncated word precedes a full word which tends to determine the stress pattern of the blend, and follows a complete word in less transparent blends. Graphemic and phonemic overlap of splinter and SW, length and currency of the former, and stress pattern of the full word, facilitate recognition. Some examples: *docuromanzo* /dokuro'mandzo/ < *documentario* /dokumen'tarjo/ [documentary]/*docu-* + *romanzo* /ro'mandzo/;<sup>9</sup> *Edilcasa* /ɛdile'kasa/ < *edile* /'ɛdile/, /e'dile/ [building trade, A]/*edilizia* /edi'littsja/ [building trade]/*-edil-* + *casa* /'kasa/; *Devitalia* /devi'talia/ < *dev-* (from Eng. *development*) + *Italia* /i'talia/; *britaliano* /brita'ljano/ < *British* /'britiʃ\*/ (GRADIT: BRITISH) + *italiano* /ita'ljano/ [Italian]; *Brangelina* /brandʒe'lina/ < *Brad (Pitt)* + *Angelina (Jolie)*, a loanword. *CLEANAP* /klin'ap/ can be seen as a less transparent blend which combines full word and splinter, if analyzed (with a bit of an effort) by the average Italian as *clean* /'klin/ + *Napoli* /na'poli/ [Naples], as against *CLEANAP* /klin'ap/ < *clean up* /klin'ap/ (cf. GRADIT: CLEAN UP) + *Napoli* /na'poli/ [Naples].<sup>10</sup>

#### 3.1.2. Clipped compounds

Ronneberger-Sibold (2006) devises one last and least transparent category, *fragment blends* or concatenations of SW beginnings in clipped compounds. Turning now to Italian, fragment blending accounts for the vast majority of subtractive word formations and is extensively used for company service and product names, also brand names and trade marks, e.g.: *Dicloream* /diclo'rewm/ < *Diclofenac sodico* /diklofe'nak 'sɔdiko/ [diclofenac sodium] + *reumatismo* /rewma'tizmo/ [rheumatism]; *Edilplast* /edi'lplast/, /'ɛdilplast/ < *edile* /'ɛdile/, /e'dile/ [building trade, A]/*edilizia* /edi'littsja/

<sup>8</sup> *Sarkozy* /sarko'zi/ and *Berlusconi* /berlu'skoni/ overlap in two attested blends, *Sarlusconi* /sarlu'skoni/ and *Berluscozy* /berlusko'zi/, which retain the stress pattern of SW<sub>2</sub> (position-based stress assignment). Though the matrix word in the contour blend is shorter than the inserted word, we suggest *Berluscozy* is a relatively more transparent coinage in that the segment /ko'zi/ has very few neighbours in Italian.

<sup>9</sup> *Vaffapensiero* can be analyzed as a blend of *vaffanculo* [fuck off, Fig: refusal] and *pensiero* [political view] or, most probably, as an endocentric compound /vaffapens'jero/, from *vaffa* /'vaffa/ (interjection; shortening, *vaffanculo*, cf. GRADIT: VAFFA) and *pensiero* /pen'sjero/.

<sup>10</sup> It is apparent that *clean up* has not gained currency nor replaced its Italian dictionary equivalents (*risanamento/pulizia/bonifica ambientale*). Indeed, although *clean up* was first attested in 1996 (GRADIT: clean up) in the sense 'ecological clean up', a simple search in the *la Repubblica Corpus (RC)* – a large, annotated corpus of newspaper Italian (380M tokens) – only returns 3 hits for ecological *clean up* and 5 hits for other types of clean-up. Further evidence for analyzing *Cleanap* as a reduction of *Clean Nap* comes from the alternating use of *CleaNap* and *CLEANAP* in the names of the corresponding webpages.

[building trade]/-*edil*- + *plastica* / plastica/ [plastic]; *Finedil* /fine'dil/ < *finanziaria*/fin-/finan'tsjarja/ [consumer credit provider] + *edile* /'ɛdile/, /e'dile/ [building trade, A]/*edilizia* /edi'littsja/ [building trade]/-*edil*-; *Findomestic* /findo'mestik/ < *finanziaria*/fin-/finan'tsjarja/ [consumer credit provider] + *domestico* /do'mestiko/ [personal/family, A].

Because fragment blends combine SW beginnings and allow closed syllables at syllable and word endings, they allow word-internal consonant cluster that are not well-formed in Italian (e.g. *lp* in *Edilplast*). Additionally, whereas in the majority of Italian non-derived nouns the stress is placed on the penultimate syllable, fragment blends do not show any regularity in this respect because truncation can occur at any point within SW<sub>1</sub> and SW<sub>2</sub>. Truncation can thus minimally affect the nucleus of the word-final syllable and retain syllable onsets of one SW (*Findomestic*, which, however, comes close to semi-complete blends such as *Finconsumo* /finkon'sumo/ < *finanziaria*/fin-/finan'tsjarja/ [consumer credit provider] + *consumo* /kon'sumo/ [consumer]), or result into significant departures from partial transparency and semi-complete blends, maximally deleting the stressed syllable in both SWs (*Dicloream*).

In other words, fragment blends show different degrees of departure from the regularities observed for complete blends, contour blends, and (to a lesser degree) semi-complete blends (cf. Sections 3.1 and Sections 3.1.1). We therefore suggest adopting the term *clipped compounds* for fragment blends and treat blends and clipped compounds as two separate though neighbouring categories within extragrammatical subtractive word-formations.

### 3.2. Morphosemantic transparency and cognitive motivation

As suggested above (cf. Section 3.1), among the factors that have an impact on the morphosemantic transparency of the blend there is the semantic plausibility of the correspondence relations (Langacker 1987: *valency relations*) between SWs (Langacker 1987: *components*) within the composite structure. In the case of subordinate endocentric blends from nouns such as *Finconsumo* (< *finanziaria*/fin- [consumer credit provider] + *consumo* [consumer]), a consumer credit provider, the correspondences are quite prominent and the blend serves both an identificatory and a categorizing function. In *L'oroschifo* [the bad horoscope], *oroschifo* (< *oroscopo* [horoscope] + *schifo* [disgust]) is used metonymically to name a live radio programme broadcasting bad horoscope predictions. In *Il gastronomo* [the well travelled gourmet], *gastronauta* (< *gastronomo* [gourmet] + *astronauta* [astronaut]) is used metonymically to name a radio programme during which Davide Paolini, also known as *il gastronomo*, provides expert information on what to eat and where to eat in Italy.

Based on analogical extensions from creative blends, the truncated word originally used in a blend may be analyzed as a word string on its own and form new word families. As such, it illustrates the productive use of symbolic constructional schemas (see Kemmer 2003 for English). Some examples in our sample are not only *-edil-* and *fin-* but also *gastro-*, a neoclassical compound, whose use is extended via *gastronauta* from specialist terminology in medicine and the life sciences to word formations such as *gastro-*

*avvertito* [gastro-aware], *gastrocolto* [gastro-expert] vs. *gastroignorante* [gastro-illiterate], *gastrocritico* [gastroreviewer], *gastrodevolution*, *gastro-follia* [unusual and bizarre dish, Ironic], *gastro-maker* [a person who decides on up-to-date gastronomic trends], etc.

Often, however, valence relations in blends as names depend on idiosyncratic features that are not easily nor immediately recognizable. This is especially the case of novel naming units from names. Coining and understanding blends from names essentially relies on encyclopaedic knowledge: inability to understand a name rather than a noun, adjective or verb is a failure in encyclopaedic knowledge of the language (Recanati 1993, in Anderson 2007: 158). Based on a highly entrenched schema (cf. Langacker 1987), male names and female names combine in English exocentric compounds naming couples (e.g., *Josalind* < *Jocelyn* + *Rosalind*, *Tyrus* < *Tyrone* + *Cyrus*, cf. Marchand 1969). Our sample comprises loan names from English that are analyzed into their constituent parts, e.g. *Brangelina* < *Brad (Pitt)* and *Angelina (Jolie)* and *Billary* < *Bill (Clinton)* + *Hillary (Rodham Clinton)*. Both represent extensions of the English Name [+MALE] – Name [+FEMALE] schema for couples and are easily accounted for within Conceptual Metaphor Theory (CMT, cf. Lakoff/Johnson 1980) and Conceptual Blending theories (Ruiz de Mendoza's 1998 Combined Input Hypothesis, Fauconnier/Turner 2002), which incorporate CMT. Its fixed, generic domains are replaced by the schematic and specific knowledge dynamically structured in the language user's partial, specific mental spaces. CB theories see word formation and the interpretation of the underlying semantic motivation in terms of the language user's online language processing: elements within contextually relevant source inputs (SIs) are activated on the basis of long-term supra- and extralinguistic knowledge, they (can) evoke common abstract schemas and images in the Generic Space (GS) and yield a Projection Space (PS) (Fauconnier and Turner 2002: blended space, or blend) which accommodates the selected elements from the SI(s), and can contain new meanings (emergent signification).

Very briefly, *Brangelina* is a reduction from a coordinate compound which extends the Name [+MALE] – Name [+FEMALE] schema for couples combining selected information from SI1 (Brad Pitt: celebrity, handsome US actor; Hollywood superstar; married to Angelina Jolie;...) and SI2 (Angelina Jolie: stunning US actress; Hollywood celebrity; married to Brad Pitt;...). Based on re-categorization from Name to Noun to Adjective via metonymy and cued inferencing (Ruiz de Mendoza 2000), in the PS the blend identifies and describes the greatly admired (and much envied) Hollywood superstar couple, thus taking on evaluative connotations. *Billary* is a slightly different example. It is based on selected idiosyncratic features epitomized by Bill Clinton and characterizing Hillary Clinton against Western cultural keywords (Williams 1973/1986) such as FAMILY, HOME (for women) and CAREER, WORK, JOB (for men) within specific schemes, frames and scenarios related to Clinton's US presidency. In the PS *Billary* can thus be interpreted as: i. a derogatory lexical blend with identificatory, descriptive and expressive meanings, which reduces an exocentric compound describing via metonymy the former First Lady and career woman Hillary Clinton; or ii. a metonymic blend based on an exocentric compound which describes a marriage which is also a joint venture.

Turning now to the Italian blends for names of individuals, applying (entrenched) schemas might work towards morphosemantic transparency and increase meaning predictability. The Name [+FEMALE] – Name [+FEMALE] schema applies to Italian female names, whereby two female names can combine in a coordinate endocentric blend (e.g. *Marilena/Malena* < *Maria* + *Maddalena*, cf. Thornton 2004a). However, based on coordinate compounds, the blends in our sample, from journalese, media language and politics metonymically combine Family names [+MALE] to denote a coalition between political parties, with LEADER OF A PARTY FOR THE POLITICAL PARTY. One example is *Berlingotti* < *Berlinguer* + *Andreotti*. Context-dependent, selected features for Berlinguer in SI1 are: leader of the major opposition party till the early eighties, working with Andreotti (leader of the Christian Democracy, the majority party) and Aldo Moro (repeatedly Prime Minister) to reach a historic compromise whereby the largest European Communist Party would have taken part in the majority coalition. Likewise, in SI2 Andreotti is the leader of the Christian Democratic Party. Based on encyclopaedic knowledge, the motivated, contextually determined metonymic reading of *Berlingotti* in PS is thus: the leaders' attempts at reaching the historic compromise and forming a coalition, or the political period at large.

Although other blends might appear to apply the schema for the 'coalition' reading, on closer inspection this reading turns out to be too context-dependent to develop into a productive constructional schema. Consider *Prodinotti*, which foregrounds opposition within a given coalition. More specifically, *Prodinotti* < *Prodi* + *Bertinotti* names Prodi's coalition government, brought down by the minority party within the coalition, led by Bertinotti. As a result, Bertinotti and his Communist Party ended up supporting the opposition coalition, led by Berlusconi (hence, *Berluscottti* < *Berlusconi* + *Bertinotti*).

Yet another example is *Sarlusconi* < *Sarkozy* + *Berlusconi/Berluscozy* < *Berlusconi* + *Sarkozy*,<sup>11</sup> where both Prime Ministers are identified by source names that are re-categorized via metonymy (NAME FOR QUALITY) to reduplicate similar features in the two prime ministers' private lives. Based on the metaphor MORE OF FORM IS MORE OF CONTENT, the blend describes and criticizes both prime ministers and their celebrity life style.

#### 4. CONCLUSIONS

The aim of this paper has been to examine Italian unfamiliar lexical blends from names and nouns, which serve an identificatory function. Although we are aware of the drawbacks of a qualitative investigation based on a restricted sample, we can now draw some preliminary conclusions. Working on the assumption that blends reach a compromise between morphotactic and morphosemantic transparency on the one hand, and the need to be memorable on the other, we placed special emphasis on the principle of saliency (cf. Dressler 1987), which turns out to motivate a number of phonological regularities better described within Optimality Theory (Bat-El/Cohen, in press) and ultimately responsible for size-based and position-based stress assignment within the blend.

---

<sup>11</sup> See Cacchiani (2010) for extensive analysis of *Billary*, *Brangelina* and *Berluscozy/Sarlusconi*.

On a gradient from more to less morphotactically transparent, we distinguish *complete blends* (*telescope blends* such as *RavennAntica* and *inclusive blends* such as *Eataly*), *contour blends*, where blend size and stress pattern are identical to the size and stress pattern of the matrix SW (*Berluscottti*, *frappuccino*), and the more peripheral subtype of *semi-complete blends*, which tend to juxtapose SW beginning and SW (*Devitalia*). Because *clipped compounds* (*Findomestic*, *Dicloream*) combine SW beginnings and show different degrees of departure from blends (especially in terms of the similarity between SW stress pattern and stress pattern of the blend), we view clipped compounds and blends as neighbouring types within the prototypical category of extragrammatical subtractive word-formations.

As far as morphosemantic transparency is concerned, although the plausibility of the semantic relation between SWs has a bearing on the semantic analyzability of the blend, processing unfamiliar naming blends from names and nouns relies on the encyclopaedic knowledge of the decoder. While, based on schemas, generalizations can only be made for combining forms that started out as blends (e.g. *fin-*, *gastro-*), one recurrent pattern was observed in Italian journalese, the language of the media ad politics, where Family names [+MALE] are integrated in lexical blends such as *Berlingotti*, *Prodinotti* or *Berluscozy/Sarlusconi*. However, the valence relations depend on idiosyncratic properties or are too context-dependent to instantiate or develop into a productive constructional schema, which ultimately reduces their meaning predictability.

## References

- ALGEO, John (1977) "Blends. A structural and systemic view." *American Speech* 52. 47-64.
- ANDERSON, John. M. (2007) *The Grammar of Names*. New York: Oxford University Press.
- BAT-EL, Outi (2006) "Blend." In: K. Brown (ed.), 66-70.
- BAT-EL, Outi/Evan-Gary COHEN (in press) "Stress in English blends: a constraint-based analysis." In: V. Renner/F. Maniez/P. Arnaud (eds).
- BAUER, Laurie (1983) *English Word-formation*. Cambridge: Cambridge University Press.
- BERTINETTO, Pier Marco (2001). "Blends and syllable structure: a four-fold comparison." In: M. Lorente/N. Alturo/E. Boix/M. R. Lloret/L. Payrató (eds), 59-112.
- BRDAR-SZABÓ, Rita/Mario BRDAR (2008) "On the marginality of lexical blending." *Jezikoslovlje* 9/1-2, 171-194.
- BROWN, Keith (ed.) (2006) *Encyclopedia of Language and Linguistics. Second edition. Volume 2*. Oxford: Elsevier.
- CUYCKENS, Hubert/Thomas BERG/René DIRVEN/Klaus-Uwe PANTHER (eds) (2003) *Motivation in Language. Studies in honour of Günter Radden*. Amsterdam/Philadelphia: John Benjamins.
- CACCHIANI, Silvia (2010) "Blends as (nick)-names – Identification, motivation, description." In: S.N. osu/G. Col/N. Garric/F. Topin (eds). 439-454.
- DE BEAUGRANDE, Robert/Wolfgang U. DRESSLER (1981) *Introduction to Text Linguistics*. London: Longman.
- DOLESCHAL, Ursula/Anna M. THORNTON (eds) (2000) *Extragrammatical and Marginal Morphology*. München: LINCOM Europa.

- DRESSLER, Wolfgang U. (1999) "What is *Natural* in Natural Morphology (NM)?" *Travaux du Cercle Linguistique de Prague* 3, 135-144.
- DRESSLER, Wolfgang U. (1987) "Word formation as part of Natural Morphology." In: W. U. Dressler/W. Mayerthaler/O. Panagl/W. U. Wurzel (eds), 99-126.
- DRESSLER, Wolfgang U. (2000) "Extragrammatical vs. marginal morphology." In: U. Doleschal/A. M. Thornton (eds), 1-9.
- DRESSLER, Wolfgang U./Willi MAYERTHALER/Oscar PANAGL/Wolfgang U. WURZEL (1987) *Leitmotifs in Natural Morphology*. Amsterdam/Philadelphia: John Benjamins.
- FAUCONNIER, Gilles/Mark TURNER (2002) *The Way we Think*. Conceptual Blending and the mind's hidden complexities. New York: Basic Books.
- GRADIT: DE MAURO, Tullio (ed.) (2003). *Grande dizionario italiano dell'uso*. Torino: UTET. [CD-rom].
- GRIES, Stefan Th. (in press) "Quantitative corpus data on blend formation: psycho- and cognitive-linguistic perspectives." In: V. Renner/F. Maniez/P. Arnaud (eds).
- GROSSMAN, Maria/Franz RAINER (eds) (2004) *La formazione delle parole in italiano*. Tübingen: Max Niemeyer Verlag.
- KEMMER, Susanne (2003) "Schemas and lexical blends." In: H. Cuyckens/T. Berg/R. Dirven/K.-U. Panther (eds), 69-97.
- KRIPKE, Saul (1981/1972) *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- LAKOFF, George/Mark JOHNSON (1980) *Metaphors We Live By*. Chicago: University of Chicago Press.
- LANGACKER, Ronald W. (1987) *Foundations of Cognitive Grammar. Volume 1. Theoretical prerequisites*. Stanford: Stanford University Press.
- LEHRER, Adrienne (2003) "Understanding trendy neologisms." *Rivista di linguistica* 15/2, 369-382.
- LÓPEZ-RÚA, Paula (2002) "On the structure of acronyms and neighbouring categories: a prototype-base account." *English Language and Linguistics* 6/1, 31-60.
- LORENTE, Mercè/Núria ALTURO/Emili BOIX/Maria-Rosa LLORET/Lluís PAYRATÓ (eds) (2001) *La gramàtica i la semàntica en l'estudi de la variació*. Barcelona: Promociones y Publicaciones Universitarias, S.A.
- LYONS, John (1977) *Semantics. Volume 2*. Cambridge: Cambridge University Press.
- MARCHAND, Hans (1969) *The Categories and Types of Present Day English Word-Formation. A synchronic-diachronic approach*. München: C.H.Beck'sche Verlagsbuchhandlung.
- MIGLIORINI, Bruno (1949) "Uso e abuso delle sigle." In: B. Migliorini (ed.), 86-90.
- MIGLIORINI, Bruno (ed.) (1949) *Conversazioni sulla lingua italiana*. Firenze: Le Monnier.
- OSU, Sylvester N./Gilles COL/Nathalie GARRIC/Fabienne TOUPIN (eds) (2010) *Constructing d'identité et processus d'identification*. Bern : Peter Lang.
- PIÑEROS, CARLOS-EDUARDO (2004) "The creation of portmanteaus in the extragrammatical morphology of Spanish." *Probus* 16, 203-240.
- RAINER, Franz/Wolfgang U. DRESSLER/Dieter KASTOVSKY/Hans C. LUSCHÜTZKY (eds) (2010) *Variation and Change in Morphology*. John Benjamins: Amsterdam/Philadelphia.
- RECANATI, François (1993) *Direct Reference. From language to thought*. Oxford: Oxford University Press.

- RENNER, Vincent/François MANIEZ/Pierre ARNAUD (eds) (in press) *Cross-disciplinary Perspectives on Lexical Blending*. Berlin/New York: Mouton de Gruyter.
- RC: *La Repubblica Corpus*. 12 August 2011. <http://sslmit.unibo.it/repubblica>.
- RONNEBERGER-SIBOLD, Elke (2006) "Lexical blends: functionally tuning the transparency of complex words." *Folia linguistica* 40/1-2, 155-181.
- RONNEBERGER-SIBOLD, Elke (2010) "Word creation. Definition – Function – Typology." In: F. Rainer/W. U. Dressler/D. Kastovsky/H. C. Luschützky (eds), 201-216.
- RUIZ DE MENDOZA IBAÑEZ, Francisco J. (1998) "On the nature of blending as a cognitive phenomenon." *Journal of Pragmatics* 30/3, 259-274.
- RUIZ DE MENDOZA IBAÑEZ, Francisco J. (2000) "The role of mappings and domains in understanding metonymy." In: A. Barcelona (ed.). 109-132.
- SPERBER, Deirdre/Dan WILSON (1986/1995) *Relevance: communication and cognition*. Oxford: Blackwell.
- TONELLI, Livia/W. U. DRESSLER (eds) (1993) *Natural Morphology: perspectives for the nineties*. Padova: Unipress.
- THORNTON, Anna M. (1993) "Italian blends." In: L. Tonelli/W. U. Dressler (eds), 143-155.
- THORNTON, Anna M. (2004a) "Formazione delle parole nell'onomastica." In: M. Grossman/F. Rainer (eds), 599-610.
- THORNTON, Anna M. (2004b) "Parole macedonia." In: M. Grossman/F. Rainer (eds), 567-571.
- WILLIAMS, Raymond (1976/1983) *Keyowrds: a vocabulary of culture ad society*. London: Fontana Press.

#### Abstract

#### ON UNFAMILIAR ITALIAN LEXICAL BLENDS FROM NAMES AND NOUNS

Italian has recently witnessed a steady increase in the use of unfamiliar lexical blends from names and nouns. While they serve an identificatory and descriptive function (in the sense of Anderson 2007), blends are created in extragrammatical morphology with careful attention to the semantic concepts encoded by the individual SWs, understanding blends from names and nouns depends on the decoder's direct or surrogate experience of the related reference.

Significantly, blends are coined out of the need to be relevant (Sperber/Wilson 1990) and show various degrees of morphotactic and morphosemantic transparency (Thornton 1986; Dressler 1987, 1999), which makes them memorable (Lehrer 2003). In this paper we therefore address blends from names and nouns within the framework of the Naturalness Theory (Thornton 1986; Dressler et al. 1987; Dressler 1999). As will be seen, although blends are not created in rule-based grammars, some overall preferences and regularities can be observed for more core items (cf. Bat-El/Cohen, in press, within the framework of Optimality Theory) under the principle of saliency (Dressler 1987).

Focusing on their morphosyntactic transparency, we provide a typology of Italian unfamiliar blends from names and nouns functioning as naming units in order to delimit the category and reassess current typologies. While we allow for a continuum of morphotactic transparency within the prototypical category of extragrammatical subtractive word-formations, we slightly adapt

Ronneberger-Sibold (2006) and suggest we distinguish between blends on the one hand and the neighbouring category of clipped compounds on the other. Moving on to morphosemantic transparency and conceptual motivation, we use concepts from Cognitive Grammar, Cognitive Metaphor Theory and theories of Conceptual Blending (cf. Lakoff/Johnson 1980; Langacker 1987; Ruiz de Mendoza 1998, 2000; Kemmer 2003) to provide some preliminary remarks on specific subtypes, and blends from personal names used in journalese and media language in particular (e.g. *Berluscotti* < *Berlusconi* + *Bertinotti*, *Berlingotti* < *Berlinguer* + *Bertinotti*). The data shows that we cannot as yet talk about instantiations or extensions of entrenched schemas.

#### Povzetek

### O MANJ POGOSTI VRSTI LEKSIKALNIH SPOJENK IZ IMEN IN SAMOSTALNIKOV

V italijanščini v zadnjem času opažamo nenehen porast rabe manj pogoste vrste leksikalnih spojenk iz lastnih imen in samostalnikov. Čeprav so spojenke v službi poimenovalne in opisne funkcije (v smislu Andersona 2007), so tvorjene znotraj nesistemske morfologije, pri čemer posebej pomembno vlogo igrata pomenska koncepta, katerih nosilka je vsaka od izhodiščnih besed. Razumevanje spojenk iz lastnih imen in samostalnikov je odvisno od razvezovalčeve neposredne ali posredne izkušnje z dano zunajjezikovno resničnostjo.

Pomembno je, da spojenke nastajajo iz potrebe po relevantnosti (Sperber/Wilson 1990) in kažejo različne stopnje morfolaktične in morfosemantčne transparentnosti (Thornton 1986; Dressler 1987, 1999), zaradi katere si jih je moč zapomniti (Lehrer 2003). V pričujočem članku se vprašanja spojenk iz imen in samostalnikov lotevamo v okviru teorije naravnosti (Thornton 1986; Dressler et al. 1987; Dressler 1999). V nadaljevanju bomo videli, da, čeprav spojenke niso tvorjene na podlagi gramatikalnih pravil, vseeno lahko v skladu z načelom jasne prepoznavnosti /ang. *salience*/ (Dressler 1987) za najbolj prototipične primere ugotovimo nekaj splošno veljavnih teženj in smernic.

Omejili smo se na morfosintaktično transparentne spojenke in podali tipologijo italijanskih manj pogostih spojenk iz imen in samostalnikov, ki delujejo kot poimenovalne enote. Naš namen je bil na eni strani zamejiti kategorijo, na drugi pa ovrednotiti že obstoječe tipologije. Čeprav se zavedamo kontinuuma morfolaktične transparentnosti znotraj prototipiske kategorije zunajsistemskih besednih krajšav, smo se rahlo oddaljili od pristopa Ronnenberger-Siboldove (2006) in predlagali, da bi razlikovali med spojenkami (ang. *blends*) na eni strani in zloženimi krni (ang. *clipped compounds*) na drugi. Na ravni morfosemantične transparentnosti in konceptualne motiviranosti se opiramo na koncepte kognitivne slovnice, kognitivne teorije metafore in teorij konceptualnega spajanja (cf. Lakoff/Johnson 1980; Langacker 1987; Ruiz de Mendoza 1998, 2000; Kemmer 2003), kar nam bo služilo za oblikovanje nekaterih uvodnih opazanj glede specifičnih podvrst in spojenk iz osebnih imen, kakršne najdemo še zlasti v časopisnem in medijskem jeziku (npr. *Berluscotti* < *Berlusconi* + *Bertinotti*, *Berlingotti* < *Berlinguer* + *Bertinotti*). Analiza pokaže, da zaenkrat še ne moremo govoriti niti o aplikaciji niti o razširitvi obstoječih shematskih vzorcev.



## THE RELATION BETWEEN PROTOTYPICAL AND MARGINAL MORPHOLOGY: THE CASE OF REDUPLICATIVE CONSTRUCTIONS

### 1 INTRODUCTORY REMARKS

This article investigates a paradigm case of a borderline phenomenon in linguistic analysis, namely constructions in which linguistic material is meaningfully iterated as well as the relation these constructions bear to prototypical and marginal areas of morphology, couching the discussion into a morphological framework along lines similar to those proposed by Zwicky/Pullum (1987) and Dressler (2000). I take the following semi-formal definition of reduplication or reduplicative construction (two interchangeably used terms in Moravcsik 1978; see also section 2) by Moravcsik (1978: 300) as a point of departure:

Utterance 1: ...A... = ...X...

Utterance 2: ...B... = ...Y...,

where A and B are non-null interpretable semantic representations that have some elements in common, X and Y are non-null syntactic, phonological, or phonetic representations, the equation sign stands for symbolic equivalence; and where Y either properly or improperly includes all of X, and a proper or improper part of X repeated n times, but B does not include a matching reduplication of A.

To this Moravcsik (1978: 300–301) adds the one a priori restriction that, within one language, the same particular meaning distinction paired off with a certain iterative form difference in utterances of the above kind has to recur in at least one other utterance pair in order to be counted as a reduplicative construction, a sensible restriction which in the present context is understood as a kind of minimal productivity requirement (essentially saying that an utterance pair thus defined should not constitute an isolated, hapax-like phenomenon in a language). Furthermore, it is important to note that in this conception reduplication results in a new, non-repetitive meaning, i.e. by repeating linguistic elements the process is not just saying the same all over again (cf. Stolz et al. 2011: 27).

As Gil (2005: 61, endnote 1; emphasis original) points out, Moravcsik's "use of the term *reduplication* is actually broader than that of most other scholars, subsuming also many constructions which would generally be considered as involving repeti-

---

\* *Author's address:* Geisteswissenschaftliche Fakultät der Karl-Franzens-Universität Graz, Institut für Sprachwissenschaft, Merangasse 70, 8010 Graz, Österreich. Email: thomas.schwaiger@uni-graz.at

tion,” thereby perfectly demonstrating the demarcational uncertainty that usually prevails in this area of research even in an oft-cited study which (in my opinion nonetheless justly) is regarded as a seminal treatment of reduplication up to this day. Moreover, additional obscurity is due to the fact that even in languages which lack reduplication in the narrower morphological sense of a productive process of inflection and/or word formation (i.e. the sense that the other scholars alluded to in Gil’s quotation above adhere to and which is also adopted in the present study) the systematic and meaningful repetition of lexical material can be considered as a widespread (I would suggest: universal; see section 4) phenomenon (cf. Stefanowitsch 2007: 29).

It is exactly this fuzzy state of affairs which the ensuing discussion tries to endow with some more clarity. Drawing on a body of earlier work as well as applying and modifying some of the more generally proposed theoretical frameworks by confrontation with a typological dataset,<sup>1</sup> the following sections tackle the questions of how to best distinguish reduplication proper (i.e. morphological reduplication) from other iteration/repetition phenomena and how to classify a newly defined concept of reduplicative constructions within a linguistic interface model differentiating between prototypical and marginal morphology.

## 2 THE RANGE OF (IR)RELEVANT ITERATION PHENOMENA

It seems more than warranted to approach the task of classifying and demarcating different iterative constructions from both a formal and a semantic angle. I deem the (interestingly in reduplication research) very frequently encountered bias towards one or the other pole of the linguistic sign to necessarily render any analysis incomplete.<sup>2</sup> The term *reduplicative construction* popularized by Moravcsik (1978) – but rather infelicitously used interchangeably with *reduplication* (a term reserved for the morphological use of repetition in this article; see section 1) by her – comes in very handy, in so far as it will be exploited here as an umbrella term for all kinds of repetitive constructions which at first sight share certain formal and functional characteristics but eventually may be further differentiated provided the right morphological model and criteria are applied (see section 3). But the starting point of inquiry must be an even more general one. Formally, phenomena ranging from the potentially infinite repetition of linguistic material on the discourse or text level right down to the duplication of single segments or sound features need to be considered at the outset. Semantically, it is indispensable to try and correlate formal properties with the functions or meanings they seem to serve

---

<sup>1</sup> The data in question (mostly) come from the Graz Database on Reduplication (*gdr*), the core of the by now expired Graz reduplication project (online at <http://reduplication.uni-graz.at/>) which was funded by the Fonds zur Förderung der wissenschaftlichen Forschung (Austrian Science Fund) and carried out during the past years under the direction of Bernhard Hurch at the Institute of Linguistics, University of Graz. The database can be directly accessed via <http://reduplication.uni-graz.at/redup/>.

<sup>2</sup> This stance largely corresponds to the sign-based approach advocated in Stolz et al. (2011).

or express, reduplicative constructions, as will become evident, being able to convey such different senses as (pragmatic) emphasis or (grammatical) plural agreement, arguably with some principle of iconicity at work as the overarching driving force (as I note only in passing with no intention to dwell upon this in itself intriguing subject any further here).

## 2.1 Excluded phenomena

The data collected and analyzed for the *gdr* within the Graz reduplication project on which this study mainly relies (see footnote 1) are a priori restricted in that certain constructions and phenomena have been excluded right from the start, the database trying to focus on morphological reduplication as far as possible (but see footnote 5). Thus, by virtue of their lack of meaningfulness, definitely (also according to Moravcsik's broader definition and use; see section 1) out of the picture are

- harmonic phenomena like assimilation and vowel harmony
  - phonological/prosodic doubling merely fulfilling templatic functions
- (cf. the definition and scope stated on the project's website and Hurch et al. 2008: 8–9)

while somewhat more arbitrarily also not included are reduplication in sign language, child language and language games.<sup>3</sup> The exclusion of sign language reduplication is only due to its incompatibility with the database structure (which was designed for integrating reduplication in oral languages) and should not be misunderstood as a neglect of its grammatical (i.e. morphological) status.<sup>4</sup> Apart from the fact that a contribution to meaning exclusively through reduplication seems to be missing in them, language games on the other hand have figured prominently in arguments for so-called expressive or extragrammatical morphology (e.g. Zwicky/Pullum 1987: 332–334), so their discussion in the present context most probably would not add anything useful to the general picture headed for (see especially sections 3.3 and 4). In essence, the last points also hold true for child language reduplication (e.g. Dressler et al. 2005) and although the question of meaning carried by reduplication in this domain seems to be somewhat less clear in many of the attested instances, some of them readily are conceivable as fulfilling – like phonological/prosodic doubling (see above) – mere templatic functions, for example.

## 2.2 Problematic phenomena

Considerably less uncontroversially, the *gdr* additionally refrains from analyzing in-depth the following phenomena, as is also explicitly stated on the Graz reduplication project's website:

---

<sup>3</sup> But see Hurch et al. (2008) for an overview of these topics next to syntactic and textual repetition as well as reduplication in phonology.

<sup>4</sup> For treatments of reduplication in German and American Sign Language see, e.g., Pfau/Steinbach (2005) and Wilbur (2005), respectively.

- repetitive syntactic operations
- recursive morphological operations
- contrastive reduplication

These arguably constitute less negligible cases in the present context. But although they will consequently be discussed in more detail right below, the following subsections will also show that these phenomena nevertheless are relatively easy to delimitate from reduplication proper in most cases.<sup>5</sup>

**2.2.1 Repetitive syntactic operations**

The repetition of elements applying across words is a discourse or syntactic phenomenon (cf. Gil 2005: 31) of either a reinforcing and/or iconic nature when looking at its content side (see the examples in Gil 2005: 39–46). Example (1) below, illustrating a form of iconic repetition (namely durativity), comes from Arapesh (Torricelli, Papua New Guinea):<sup>6</sup>

(1) ... <i>n-a-uli</i>	<i>nobag, n-a-na</i>	<i>n-a-na</i>
3SG.M SUBJ-R-hunt	dogs 3SG.M SUBJ-R-go	3SG.M SUBJ-R-go
 <i>n-a-nak</i>	 <i>n-a-nú</i>	 <i>nobag.</i>
3SG.M SUBJ-R-go	3SG.M SUBJ-R-with	dogs
‘... he went hunting with dogs and went and went and went a long way with the dogs.’		
(Conrad/Wogiga 1991: 53)		

Even more clear-cut are constructions that (optionally or obligatorily) use a linker explicitly conjoining repeated elements (e.g. the conjunction *and* in the English translation of the Arapesh example above) or that offer prosodic cues pointing at the relative independence of their component parts (e.g. pauses between the elements which are repeated, clearly separate intonational curves for each repeated unit, etc.). In most cases the size of the repeated constituent and/or the number of its (potential) repetitions (i.e. essentially formal characteristics) make clear whether one is dealing with repetition or reduplication. These and other diagnostics are especially fruitful for the distinction in languages that abundantly have both construction types.<sup>7</sup> Of

---

<sup>5</sup> At this point it should be noted that, owing to independent research interests, the *gdr* also includes onomatopoeic expressions, ideophones and lexical reduplications, i.e. superficially reduplication-like constructions without an independently existing simplex form. For obvious reasons these do not figure in this study.

<sup>6</sup> The following glossing conventions are used in this article: 3 = 3<sup>rd</sup> person, ADJ = adjective, F = feminine, M = masculine, NMLZ = nominalizer, PL = plural, R = realis, SG = singular, SUBJ = subject.

<sup>7</sup> For a comprehensive treatment of such diagnostics and their employment in a single language (Riau Indonesian) see Gil (2005).

course, on the one hand, the extremely rare cases of grammatical triplication – e.g. Mokilese *roar* ‘to give a shudder’ – *roar-roar* ‘to be shuddering’ – *roar-roar-roar* ‘to continue to shudder’ (Harrison 1973: 426) – could then be mistaken for syntactic repetition judging superficially from the number of repetitions alone. But for the Mokilese example this can easily be rebutted by the triplication’s systematic meaning opposition vis-à-vis its simplex and properly reduplicated counterparts as well as the inability to add more than two copies to the base for the same result. A language like English, on the other hand, can pose problems if isolated examples like *he is very very bright* (Moravcsik 1978: 301) are investigated. It is maintained here that this does not instantiate a case of reduplication (English most certainly being a non-reduplicating language) but rather repetition, a view which is supported by Moravcsik (1978: 312) herself when she states that “it is perhaps true in all languages that an emphatic modifier [...] can be open-endedly reduplicated [i.e. repeated; see above; TS] for additional degrees of emphasis.” This alleged open-endedness of the process runs counter to the usual restrictions found with morphological reduplication.

### 2.2.2 Recursive morphological operations

Having thrown a glance at the far outskirts of morphology, paying special attention to repetition phenomena in discourse and syntax in the preceding section, this one touches upon a genuinely morphological operation sometimes confusable with reduplication. A German word like *Ur-ur-großvater* ‘great-great-grandfather’ illustrates what could well be mistaken for an instance of affix reduplication exemplified here by Fijian (Austronesian, Fiji): *vanua* ‘country’ – *vēi-vanua* ‘various countries’ – *vēi~vēi-vanua* ‘larger number of countries’ (Schütz 1985: 367). While in Fijian the collective or distributive prefix *vēi-* is reduplicated for the expression of greater number in the noun, the multiple prefixation of *Ur-* in German is a recursive operation in which the (potentially infinite) repetition of the prefix is accompanied by simultaneous addition of the same meaning with each single occurrence. It follows that the meaning of *Ur-ur-ur-ur-ur-großvater* ‘great-great-great-great-great grandfather’, for example, can be arrived at by knowledge of the basic meanings of *Ur-* and *Großvater* and simply adding them up, thus almost mechanically calculating the intended overall linguistic content of the resulting form. Every instance of *Ur-* exerts its own independent meaning contribution upon the respective base words (i.e. *Großvater*, *Ur-Großvater*, *Ur-ur-großvater*, etc.) while the reduplication of *vēi-* above is a non-repetitive (see section 1), grammaticalized expression for a special plural meaning (plus it is restricted to the duplication of the affix so that, say, *\*vēi~vēi~vēi-vanua* would be ungrammatical). Furthermore, this stacking of affixes in recursive morphological operations most probably only works with certain derivational meanings lending themselves easily to multiplication, a point which in contrast is not met by the rather inflectional affix reduplication example from Fijian above.

### 2.2.3 Contrastive reduplication

The last of the problematic cases excluded by the *gdr* treated in this section is at first sight perhaps also the hardest to keep apart from reduplication proper. Contrastive reduplication (or contrastive focus reduplication) can be illustrated by (colloquial) English *SALAD-salad* (as opposed to, e.g., *tuna salad*) which yields a sort of prototypical reading of the reduplicated item in question (see Ghomeshi et al. 2004). Formally, it looks like reduplication in that it is restricted to one repetition of the base, but it is quite different from it in that also whole phrases as in *I don't LIKE-HIM-like-him* (Ghomeshi et al. 2004: 321) may be duplicated. Functionally, it seems to be restricted to very special contexts in which a contrast to a (explicitly or implicitly) present counterpart wants to be achieved. Nevertheless, a certain affinity to genuine reduplication meanings like specificity ('specific type of X'; see Kiyomi 1995: 1154) cannot be denied. But given the construction's special information structural and formal status (Ghomeshi et al. 2004: 342 sum up that "the base is of variable size, in both syntax and phonology"), all in all the exclusion of contrastive focus reduplication from the *gdr* is not at odds with the overall morphological conception of the Graz database.

### 2.3 Residual phenomena

What has been done in this section was mainly to recapitulate and at points adapt existing viewpoints of rather well investigated areas related to the adequate delimitation of linguistic iteration phenomena. Looking through the eyes of morphology, phenomena from right across its external boundaries (repetitive syntactic operations; see section 2.2.1), other morphological operations (namely recursive ones; see section 2.2.2) and constructions with an unclear status as to their appropriate linguistic module of handling (contrastive reduplication; see section 2.2.3), all showing considerable similarities with morphological reduplication (the main focus of this article), have been examined.

The progress that this study aims to make lies in concentrating on the residue that is left (i.e. reduplication in the narrow sense) and demonstrating – by looking at specific typological data from the *gdr* – that also inside this domain more fine-grained, essentially morphology *internal* distinctions can and have to be made, as not everything which looks like reduplication actually is located on the same level, a lesson which in their own way the preceding subsections have already taught from far less surprising perspectives.

## 3 REDUPLICATIVE CONSTRUCTIONS

In the beginning of section 2 it was already announced that the term *reduplicative construction* will be furnished with a new reading for the purposes of this article. Generally speaking, it is here understood as an umbrella term for constructions exemplified by (2) and (3) as well as by (4) to (6) below, each representative being taken from the *gdr*:

- (2) Eastern Oromo (Afro-Asiatic, Ethiopia)  
*isáan did-dikk'oo*<sup>8</sup>  
 they PL~small-ADJ.F  
 'they are small' (Owens 1985: 93)
- (3) Chamorro (Austronesian, Guam)  
*Hu li 'e ' i há-hatsa.*  
 I see the NMLZ~lift  
 'I saw the one that was lifting.' (Topping 1973: 102)
- (4) Lavukaleve (East Papuan, Russell Islands)  
*leon* 'quickly' – *le-leon/leo-leon/leon-leon* 'hurry up!' (Terrill 2003: 36)
- (5) Swahili (Niger-Congo, Tanzania)  
*-chapua* 'speed up' – *chapu-chapu* 'hurry up!, quick!' (Novotna 2000: 60)
- (6) *Tukang Besi* (Austronesian, Indonesia)  
*koruo* 'many' – *kor-u-'uo*<sup>9</sup> 'certainly many'  
  
*amai* 'they' – *am-a-'ai* 'certainly them' (Donohue 1999: 42)

On formal grounds it is fairly easy to distinguish all of these examples from the phenomena discussed in section 2.2: Most of them constitute instances of partial reduplication (a quite safe hint at the word-internal, i.e. morphological, status of the process; see Gil 2005: 31) and all of them show exactly one repetition of their respective bases. Semantically, (2) and (3) are equally unproblematic as the formally partial reduplication pattern is accompanied by the unambiguously grammatical functions of plural agreement (an inflectional category) and word-class derivation, respectively. On the content side things are not that easy with (4), (5) and (6), however. This shall be examined more closely in what follows.

### 3.1 Reduplicative imperatives

In a recent typological survey of imperatives in language one comes across the statement that “reduplication is never the sole marker of an imperative” (Aikhenvald 2010: 33). This falls in line with an even more recently compiled list of inflectional categories not expressed by total reduplication which inter alia displays the imperative category

---

<sup>8</sup> An adjective pluralized for agreement by reduplication takes the feminine adjectival class marker suffix (see Owens 1985: 93).

<sup>9</sup> For details on this form of vowel reduplication, including the presence of the glottal stop (ʔ), see section 3.2 including footnote 11.

(see Stolz et al. 2011: 194). But both of these assertions from the pertinent typological and reduplication literature are obviously in conflict with examples (4) and (5) above.<sup>10</sup>

Admittedly, neither for Lavukaleve nor for Swahili do the respective authors explicitly speak of reduplication as expressing the imperative, the reader can merely infer this from the translations of the examples given. Still it seems to be the most natural reading as the Lavukaleve example (4) appears in the subsection on reduplication which is part of the higher-level discussion of morphophonemic processes in the language in general (see Terrill 2003: 35–36) while the Swahili example (5) is adduced by Novotna (2000: 60) when she speaks about the reduplication of verbs which lose their final vowel after having explicitly stated before that reduplication is understood as a fully grammaticalized process serving either morphosyntactic or word-formation purposes (cf. Novotna 2000: 58).

Nevertheless, this section does not intend to falsify Aikhenvald’s and Stolz et al.’s typological claims in their entirety. Neither does it want to establish the imperative as a legitimate reduplicative category on a par with much more frequent meanings like (nominal or verbal) plurality, diminution or intensification (see Kiyomi 1995; for an example of intensifying reduplication see section 3.3). Rather it tries to pave the way for a motivated view in which the latter, prototypical reduplication meanings are adequately distinguished from marginal meanings like the imperative which in turn need to be kept apart from the repetition phenomena encountered in section 2.2. This calls for a gradient theoretical device appropriately capturing the facts at hand which with the highest probability seems to be hidden in a differentiated view of morphology itself. But first a closer look at a further construction type from the above examples is required.

### 3.2 Pragmatic emphasis

What remains to be discussed with example (6) is another special case of reduplication found in the *gdr*. Again it is the author of the primary source himself who discusses and classifies the phenomenon in question jointly with other much more clearly grammatical cases of the reduplication process (see Donohue 1999: 42). At the same time, though, its special status is quite vividly uncovered by describing (stressed) vowel reduplication<sup>11</sup> in *Tukang Besi* as being “used for pragmatic effect to emphasise the truth values of one word sentences” (Donohue 1999: 42). These quite severe formal and functional restrictions cast serious doubt upon the process in question as constituting an instance of what is commonly called morphological reduplication. Next to imperatives (see section 3.1) this is a second argument for a differentiated view of morphology in general and morphological reduplication in particular, finally outlined in the next subsection.

---

<sup>10</sup> The fact that the Swahili simplex verb in (5) loses its final vowel in the course of reduplication (see Novotna 2000: 60) should not distract from the fact that the resulting structure is one of total reduplication.

<sup>11</sup> The occurrence of the glottal stop in (6) is non-phonemic: It is inserted between two vowels that occur in adjacent syllables as a result of morphological, syntactic or pragmatic factors in *Tukang Besi* (cf. Donohue 1999: 25).



### 3.3 Reduplicative constructions vis-à-vis prototypical and marginal morphology

We are now able to pinpoint what actually is understood by the concept of a reduplicative construction in this article: The term basically refers to conventional morphological reduplication but at the same time is supplemented by components of theoretical frameworks stressing the non-uniformity of morphology and morphological processes as envisaged in the programmatic outlines of Zwicky/Pullum (1987) and Dressler (2000). Essentially conflating these two approaches, the basic division drawn here is one of prototypical (or core) morphology versus marginal (or peripheral or – and this is crucial – expressive) morphology. The main deviation from tradition that I introduce thus is a consequence of granting the marginal areas of morphology (i.e. less productive phenomena transgressing external and internal boundaries of morphology) a certain amount of expressivity while when looking at the basic works mentioned above Dressler's extragrammaticality (i.e. that which lies outside of grammar) and Zwicky/Pullum's expressive morphology more or less can be equated (cf. Dressler 2000: 1).

Applied to reduplication, such a view predicts that prototypical and marginal instances – which collectively will be referred to as reduplicative constructions – are possible for this specific morphological process as well, a prediction borne out by the *gdr* data, recall the inflectional and derivational examples (2) and (3) as well as the special formations (4), (5) and (6). It is argued that the latter three are located at the expressive margin of morphology, slowly passing over to non-morphological types of repetition. This needs some elaboration.

Sections 3.1 and 3.2 have to some extent dwelled upon reduplicative imperatives and *Tukang Besi* vowel reduplication, respectively. It has been shown that the pragmatic nature of the *Tukang Besi* case is first-handedly described in Donohue (1999). Prototypically performing directive speech acts in attempting to make the addressee do something (cf. Aikhenvald 2010: 14, endnote 1), imperatives, too, are a very pragmatically oriented category of language. And as Dressler (2000: 7) points out in his plea for the recognition of marginal morphology, a “close relation to pragmatics, i.e., morphopragmatics, is non-prototypical for morphology”.<sup>12</sup> Furthermore, one can test the two constructions under scrutiny against the seven criteria for expressive morphology assembled by Zwicky/Pullum (1987: 335–338):

The *pragmatic effect* of both construction types has already repeatedly been pointed out, the concrete imperative examples (4) and (5) even seeming to contain additional emphasis arguably due to some kind of time pressure present in situations in which they are normally uttered. Concerning the *promiscuity with regard to input category* it is interesting to note that not only verbs (i.e. the expected locus of imperative marking) are attested as bases, see (4), while when judging from the description and examples of *Tukang Besi* vowel reduplication the input category of this latter process

---

<sup>12</sup> In this context also note the often minimal or zero overt expression (i.e. lack of surface morphology) of the imperative category in languages, e.g. German or Turkish, which use the bare verbal stem to this end.

should be potentially open. As demonstrated in (4), *alternative outputs* definitely exist in Lavukaleve, the Tukang Besi case in contrast is quite restricted formally. *Special syntax* obviously is present in Tukang Besi vowel reduplication due to its exclusive occurrence in one word sentences, but while syntactic make-up also plays an important role in the typology of imperatives (see Aikhenvald 2010: 92–97), no information for the Lavukaleve and Swahili examples is available on this topic. In addition, no useful information for either construction type can be extracted from the relevant sources for the criteria *promiscuity with regard to input basehood*, *imperfect control* and *interspeaker variation*. From a quick glance, all in all this may not leave a very satisfying impression in the first instance. But as the last three criteria do not explicitly fail (the relevant information is simply lacking) and three others are definitely fulfilled for each construction, it does not seem too daring to ascribe them some expressivity and thus a special status within reduplicative morphology after all. This gains additional support from the requirement formulated by Zwicky/Pullum (1987: 338; emphasis original) that “for a phenomenon to be classified as expressive morphology, it *must* have a significant number of the above criterial properties” which in itself is very vague and open to interpretation.

Remember that especially discourse or syntactic repetition has been characterized by the function of reinforcement (see section 2.2.1). In the conception put forward here it is not regarded as a mere coincidence that morphology external repetition and morphology internal intensifying reduplication (e.g. Arapesh *ripok* ‘cut’ – *ri~ripok* ‘hack up’; Dobrin 2001: 36) should eventually meet and yield boundary phenomena like the ones described in detail above. Some time has been devoted there precisely to show the expressive (or reinforcing) character such formations convey. On the other hand one has to acknowledge their formal restrictions which can go as far as constituting a single vowel copying pattern as the one found in Tukang Besi. Considering the constructions’ form and meaning properties they consequently neither are repetition nor prototypical reduplication but rather a boundary phenomenon. Contra Zwicky/Pullum (1987: 338) this surely is a continuum approach to the morphological component perfectly in line with the conclusion drawn in Dressler (2000). For I am convinced that if adopted and steadily tested against an ever-growing amount of relevant data, the proposed model of prototypical and marginal morphology in the realm of reduplication can eventually “be elevated to the rank of a systematic study of basic questions which are likely to illuminate research in morphology at large” (Dressler 2000: 8).

#### **4 CONCLUSION: THE EMERGING GENERAL PICTURE OF REPETITION PHENOMENA IN LANGUAGE**

Returning to the broad starting point of this study outlined in section 2, one can now try and see how the discussion at hand in general bears on the classification and possible development of repetition phenomena in language. What emerges (for details see below) is almost the exact mirror-image of the summarized prototype account (the details of which need not concern us here) by Stolz et al. (2011: 69). This is hardly

surprising since these authors concentrate exclusively on formal aspects in their work when they differentiate between exact and non-exact partial and total reduplication in syntax and morphology from the point of view of prototypicality, with exact syntactic total reduplication (note the again much wider use of the term reduplication as devised here) constituting the prototype. This work, in contrast, concentrates on form-meaning pairings having the special property of displaying repetitive elements (in a fairly wide sense at first) and their gradual narrowing down from unrestricted reinforcing repetition to constrained reduplication (including partial reduplication but by no means confined to it) for grammatical purposes. I already tried to show how marginal morphological reduplicative constructions fit into this picture.

A promising way to comprehensively capture meaningful repetition phenomena ranging from discourse to morphology is provided by grammaticalization theory, the reduction of multiple repetition to simple binary reduplication structures being an essential feature of the grammaticalization process (cf. Dressler et al. 2005: 456). Gradual transitions and hence boundary, marginal and prototypical grammatical phenomena fall out naturally from such a conception and enable a more adequate take on the reduplication process and its related constructions from a typological perspective.

A little *gedankenexperiment* to conclude this study: In the colloquial style of my native language (Austrian) German it seems perfectly fine to say something like *schnell schnell* ‘quick(ly) quick(ly)’ (but also more repetitions like *schnell schnell schnell*, *schnell schnell schnell schnell*, etc. are possible) approximately meaning ‘quickly!, do it quickly!, hurry up!’. This is very reminiscent of (one is tempted to say: parallel to) the Lavukaleve example in (4). Now, German is by all means as non-reduplicating as English (cf. Section 2.2.1), still it has a reduplicative construction very similar to one found in a quite extensively reduplicating language. Of course I do not claim that German is special in this regard. Although not a native speaker, I can very well imagine a similar construction being possible in English, for example. I would even go as far as conjecturing that it is a universally available construction type that we are dealing with here. But recall two facts: First, in the German case multiple repetitions are possible. Second, Lavukaleve shows multiple patterns for the same meaning ranging from total reduplication to partial CV-reduplication. In accordance with any garden-variety theory of grammaticalization such formal differences in restriction and reduction are indicative correlates assigning German a less grammaticalized status concerning the construction under scrutiny when compared to its Lavukaleve correspondent. All this is expected under the view purported here. The universal availability of repetitive imperatives suggested above is linked to the universality of general reinforcing repetition in discourse, the difference being that in German the restricting systematization has not progressed as far as in Lavukaleve, making the repetitive imperative a morphology external phenomenon in the former and a marginal morphological (i.e. a truly reduplicative) phenomenon in the latter language. It is not entrenched deeper in the reduplicative system of Lavukaleve as has been shown in section 3.3. The question is of course if further entrenchment is a possibility at all without showing any additional formal markers (e.g. imperative particles, simultane-

ous imperative affixes, etc.). Aikhenvald's and Stolz et al.'s typological generalizations (see section 3.1) suggest a negative answer to this question. But although it will not be possible to settle the issue here once and for all, it needs to be stressed that the findings made in this article at least render a total exclusion of reduplicated-only imperatives from the domain of morphological reduplication untenable. The fact that they nonetheless are to a certain extent different from constructions normally called reduplication merely highlights the necessity of a more fine-grained conception of morphology and its relation to neighboring linguistic modules.

## Bibliography

### Primary sources/corpus

*Graz Database on Reduplication (gdr)*. 9 July 2011. <http://reduplication.uni-graz.at/redup/>.

### References

- AIKHENVALD, Alexandra Y. (2010) *Imperatives and commands*. Oxford: Oxford University Press.
- AMMANN, Andreas/Aina URDZE (eds) (2007) *Wiederholung, Parallelismus, Reduplikation: Strategien der multiplen Strukturanwendung*. Bochum: Universitätsverlag Dr. N. Brockmeyer.
- CONRAD, Robert J./Kepas WOGIGA (1991) *An outline of Bukiyip grammar*. Canberra: The Australian National University.
- DOBRIN, Lise M. (2001) "Arapesh." In J. Garry/C. Rubino (eds), 33–38.
- DOLESCHAL, Ursula/Anna M. THORNTON (eds) (2000) *Extragrammatical and marginal morphology*. München: Lincom Europa.
- DONOHUE, Mark (1999) *A grammar of Tukang Besi*. Berlin/New York: Mouton de Gruyter.
- DRESSLER, Wolfgang U. (2000) "Extragrammatical vs. marginal morphology." In: U. Doleschal/A. M. Thornton (eds), 1–10.
- DRESSLER, Wolfgang U./Katarzyna DZIUBALSKA-KOŁACZYK/Natalia GAGARINA/Marianne KILANI-SCHOCH (2005) "Reduplication in child language." In: B. Hurch (ed), 455–474.
- GARRY, Jane/Carl RUBINO (eds) (2001) *Facts about the world's languages: An encyclopedia of the world's major languages: Past and present*. New York/Dublin: H. W. Wilson Press.
- GHOMESHI, Jila/Ray JACKENDOFF/Nicole ROSEN/Kevin RUSSELL (2004) "Contrastive focus reduplication in English (The salad-salad paper)." *Natural Language & Linguistic Theory* 22/2, 307–357.
- GIL, David (2005) "From repetition to reduplication in Riau Indonesian." In: B. Hurch (ed), 31–64.
- GREENBERG, Joseph H. (ed) (1978) *Universals of human language*. Vol. 3. Stanford: Stanford University Press.
- HARRISON, Sheldon P. (1973) "Reduplication in Micronesian languages." *Oceanic Linguistics* 12/1–2, 407–454.
- HURCH, Bernhard (ed) (2005) *Studies on reduplication*. Berlin/New York: Mouton de Gruyter.

- HURCH, Bernhard/Motomi KAJITANI/Veronika MATTES/Ursula STANGEL/Ralf VOLL-MANN (2008) *Other reduplication phenomena*. 9 July 2011. [http://reduplication.uni-graz.at/texte/Other\\_Red\\_Phen.pdf](http://reduplication.uni-graz.at/texte/Other_Red_Phen.pdf).
- KIYOMI, Setsuko (1995) "A new approach to reduplication: A semantic study of noun and verb reduplication in the Malayo-Polynesian languages." *Linguistics* 33/6, 1145–1167.
- MORAVCSIK, Edith A. (1978) "Reduplicative constructions." In: J. H. Greenberg (ed), 297–334.
- NOVOTNA, Jana (2000) "Reduplication in Swahili." *Afrikanische Arbeitspapiere* 64, 57–73.
- OWENS, Jonathan (1985) *A grammar of Harar Oromo (Northeastern Ethiopia)*. Hamburg: Helmut Buske Verlag.
- PFAU, Roland/Markus STEINBACH (2005) "Backward and sideward reduplication in German Sign Language." In: B. Hurch (ed), 569–594.
- SCHÜTZ, Albert J. (1985) *The Fijian language*. Honolulu: University of Hawaii Press.
- STEFANOWITSCH, Anatol (2007) "Wortwiederholung im Englischen und Deutschen: Eine korpuslinguistische Annäherung." In A. Ammann/A. Urdze (eds), 29–45.
- STOLZ, Thomas/Cornelia STROH/Aina URDZE (2011) *Total reduplication: The areal linguistics of a potential universal*. Berlin: Akademie Verlag.
- TERRILL, Angela (2003) *A grammar of Lavukaleve*. Berlin/New York: Mouton de Gruyter.
- TOPPING, Donald (1973) *Chamorro reference grammar*. Honolulu: University of Hawaii Press.
- WILBUR, Ronnie B. (2005) "A reanalysis of reduplication in American Sign Language." In: B. Hurch (ed), 595–623.
- ZWICKY, Arnold M./Geoffrey K. PULLUM (1987) "Plain morphology and expressive morphology." *Berkeley Linguistics Society* 13, 330–340.

#### Abstract

#### THE RELATION BETWEEN PROTOTYPICAL AND MARGINAL MORPHOLOGY: THE CASE OF REDUPLICATIVE CONSTRUCTIONS

The article investigates a paradigm case of a borderline phenomenon in linguistic analysis: constructions in which linguistic material is meaningfully iterated (or repeated) and their relation to prototypical and marginal areas of morphology. The fuzzy state of affairs prevailing in this research area is described and a survey of relevant and irrelevant iteration phenomena is undertaken. The discussion finally narrowing in on morphological reduplication, the data dealt with (mostly) come from the typologically oriented Graz Database on Reduplication (*gdr*). In light of certain data encountered there (i.e. reduplicative imperatives and a pragmatically emphatic vowel copying construction), the morphological process of reduplication is further differentiated, the general term reduplicative construction (Moravcsik 1978) thereby being endowed with a special meaning subsuming both prototypical and marginal instances of the process under scrutiny, couching the investigation into a morphological framework along lines similar to those proposed by Zwicky/Pullum (1987) and Dressler (2000). The study concludes with a tentative general picture of repetition phenomena in language in which grammaticalization theory can comprehensively account for such constructions that linguistically range from discourse to morphology. Finally, the potential benefit of the present approach for any typological undertaking in the realm of the reduplication process is highlighted.

RAZMERJE MED PROTOTIPSKO IN MARGINALNO MORFOLOGIJO: PRIMER  
REDUPLIKATIVNIH KONSTRUKCIJ

Članek raziskuje paradigmatičen primer mejnega pojava v jezikoslovni analizi: gre za konstrukcije, kjer je jezikovno gradivo smiselno podvojeno (oziroma ponovljeno), in odnos teh konstrukcij do prototipskih in marginalnih področij morfologije. Opisujemo nejasnost, v katero je potopljeno omenjeno raziskovalno področje, in se lotimo pregleda relevantnih in nerelevantnih podvojitvenih pojavov. Razprava se nato omeji na morfološko reduplikacijo. Gradivo, ki služi za analizo, prihaja večinoma iz tipološko zamejene graške podatkovne baze reduplikacij (Graz Database on Reduplication, *gdr*). V luči nekaterih pridobljenih podatkov (npr. reduplikativni velelniki in pragmatično pogojene emfatične konstrukcije s ponovitvijo samoglasnikov) je morfološki proces reduplikacije še dodatno razdelan, tako da je splošnemu terminu reduplikativnih konstrukcij (Moravcsik 1978) pripisan poseben pomen, ki pokriva tako prototipske kot marginalne primere procesa, ki ga raziskujemo, in ki bo raziskavo umestil v morfološki okvir, ki je blizu pristopom Zwickyja in Pulluma (Zwicky/Pullum 1987) in Dresslerja (Dressler 2000). Študija se zaključuje s poskusom zarisa splošne slike ponovitvenih pojavov v jeziku, s katerim lahko teorija gramatikalizacije celovito razloži konstrukcije, ki jih sicer jezikoslovje umešča med diskurz in morfologijo. Poudarimo tudi pomen, ki ga ima lahko opisani pristop za vse nadaljnje poskuse izdelave tipologij na področju reduplikativnih procesov.

## NOMINAL DIMINUTIVES IN SLOVENE AND ENGLISH

The present article deals with diminutiveness in Slovene and English nouns within the frameworks of semantics and word-formation; the focus is on word-formational and syntactic characteristics of diminutive nominal derivatives in the two languages concerned. On grounds of language-typological differences between English (analytic language) and Slovene (synthetic language), diminutiveness as language phenomenon is often expressed differently in these two languages: in Slovene, diminutiveness is primarily expressed within one single lexeme, while in English it is primarily expressed syntactically, extending over the boundaries of one single lexeme. In both languages, however, regardless of the either lexemic (word-formational) or syntactic aspect of diminutiveness, the semantic aspect also needs to be taken into consideration, i.e. denotative and connotative diminutiveness as well as diminutiveness in the role of terminological metaphor.

### 0 INTRODUCTION

A contrastive analysis of nominal diminutives used in modern Slovene and English clearly shows that diminutive formation and use of diminutives in Slovene remain tied to the morphological characteristics of nouns and, consequently, their morphological-lexemic features. On the other hand, the focus of diminutive formation and use in English remains bound to the syntactic use, or rather, the respective syntactico-semantic use of a given lexeme; however, the context also needs to be taken into consideration when analyzing and contrasting expressions of diminution. In all languages, diminutiveness is a basic meaning-forming element, which can, however, be realized predominantly morphologically, as is the case in Slovene, or predominantly syntactically, as is the case in English. As semantic expansion diminutiveness also plays a crucial role in the development of terminology – in this case the diminutive as language metaphor gains semantic independence and becomes a technical term.

In the following analysis, ways of expressing diminutiveness in Slovene and in English will be contrasted. While in the Slovene language diminutiveness is primarily expressed by means of suffixal endings within a lexeme, the English language

---

\* *Author's address:* Filozofska fakulteta, Oddelek za anglistiko in amerikanistiko, Aškerčeva 2, Ljubljana, Slovenia. Email: eva.sicherl@ff.uni-lj.si

\*\* *Author's address:* Filozofska fakulteta, Oddelek za slovenistiko, Aškarčeva 2, Ljubljana, Slovenia. Email: andreja.zele@zrc-sazu.si

only occasionally expresses diminutiveness in this same, synthetic, way. More often, expressing diminutiveness in English demands going beyond the borders of a lexeme, stretching over an entire syntactic structure or even sentence.

## 1 DIMINUTIVENESS AND CONNOTATION

Basically, two types of nominal diminutive can be distinguished in both languages: diminutives with the denotative meaning of ‘smallness’, e.g.

*gumbek* [small button],  
*račka* [little duck or duckling],

and diminutive nouns with the predominant connotative or expressive shade of meaning, used by speakers when they intentionally modify the object,<sup>1</sup> e.g.

*advokatek* [unimportant lawyer],  
*revček* [helpless person],  
*ženička* [helpless, unimportant woman],  
*člančič* [worthless, unimportant article].

The expressive value of connotation becomes even clearer when used in the context:

*Takle advokatek ti lahko naredi veliko škode!*  
*My second telephone call produced a lecturette on the regulations governing the use of Elastoplast in hotels.* (Guardian 1995, quoted in Adams 2001: 13)

However, in most cases of nominal diminutives, a double semantic value can be observed: apart from the inherent denotative meaning, practically always some connotative meaning is expressed too. Denotation and connotation usually overlap within one and the same diminutive lexeme, which accounts for blurred boundaries between denotation and connotation.<sup>2</sup>

Two other important elements also play a role in expressiveness, namely datedness and standardness or (in)formality of the diminutive, as can be observed in the following Slovene examples, derived from two bases *dekle* and *punca* [girl]<sup>3</sup>:

---

<sup>1</sup> The theory of expressiveness by the Czech linguist Zima (1961), distinguishing between *inherent expressiveness* (i.e. expressiveness in the lexical meaning) and *adherent expressiveness* (i.e. expressiveness acquired from the text), was first introduced to Slovene linguistics by Vidovič Muha (1971/72, 1972).

<sup>2</sup> More on this aspect of diminutiveness in Slovene is to be found in Vidovič Muha (2000: 97 ff).

<sup>3</sup> Markedness of content is indicated in the *Slovar slovenskega knjižnega jezika* (1970-1991) by the following labels: *ekspr(esivno)* [expressive], *evfem(ično)* [euphemistic], *iron(ično)* [ironic], *ljubk(ovalno)* [endearing], *slabš(alno)* [pejorative], *šalj(ivo)* [facetious], *vulg(arno)* [vulgar] and/or temporal-frequency labels such as *raba narašča* [frequent], *raba peša* [rare], *star(insko)* [archaic], *zastar(elo)* [dated]. Vidovič Muha (1971/72) specifically comments on lexemes with combined labels expressive and standard as well as expressive and colloquial.



*dekle* > *dekletce/deklīca* (standard)  
*deklič* > *dekliček* (dated)  
*deklinica* (dialect)  
*punca* > *punčī/punčīca/punčka* (standard)  
*punče* (n) / *punček* (m) (dated)  
*punčīka* (dated)

The examples listed above illustrate the immense word-formational potential of the Slovene language in this respect; it seems that English needs to compensate for its lack of similar synthetic means by employing other, more analytic (syntactic) methods of expression (e.g. *Ta punčka že kaže talent.* – *This little girl is already showing talent.*).

In the present analysis, diminutiveness, and along with it frequent expressiveness, is limited to the framework of word-formation and semantics of Slovene nouns, or, rather, Slovene (suffixal) derivatives, contrasted with their English equivalent expressions in specific syntactic use. The starting point is therefore the Slovene diminutive, compared with its English parallel expression; however, these English equivalents mostly have to be expressed syntactically and not word-formationally due to the specifics of the English language (see also Klinar 1996).

## 2 SEMANTIC AND WORD-FORMATIONAL PATTERNS FOR EXPRESSING DIMINUTION IN SLOVENE (WITH THEIR ENGLISH EQUIVALENTS)

In Slovene the following endings are primarily used to express diminutiveness: *-ek* (*gumbek* [small button]), *-ica* (*mizica* [low table], *kozica* [kid] - [goat], *babica* [granny]), *-ika* (*mucika* [kitten]), *-ka* (*nitka* [thread]), *-ca* (*klopca* [bank]), *-ko* (*dečko* [little boy]), *-ec* (*bratec* [baby brother]), *-e* (*otročje* [kid - child]), *-ce* (*mestece* [small town]), *-ič* (*tatič* [little thief]), *-i* (*babi* [granny]) (prim. Bajec 1950, Toporišič 1976, Vidovič Muha 1995).<sup>4</sup> In Slovene, all nominal suffixal derivatives, regardless of the original declension-type of their base, enter the first nominal declension pattern of their respective gender, the only exception being the feminine diminutives ending in *-i* (e.g. *mami*, *babi*).

Qualitative adjectives, or, more precisely, qualitative adjectives with a positive emotional value ('nice, pleasant') or a negative emotional value ('unpleasant') and the dimensional spatial adjective *majhen* [small] or the temporal adjective *mlad* [young] are transformed into a suffixal ending<sup>5</sup>:

otrok-ec ← [prijeten] otrok[-ø], []→-ec, otrok-  
profesor-ček ← [neprijeten] profesor[-ø], []→-ček, profesor-  
grič-ek ← [majhen] grič[-ø], []→-ek, grič-  
medvedek ← [mlad] medved[-ø], []→-ek, medved-

<sup>4</sup> The productivity and frequency of these suffixes have been commented on in the word-formational studies by Bajec (1950: 125–126) and Stramljič Breznik (2007, 2010).

<sup>5</sup> See Vidovič Muha (1995: 158).

The English language only occasionally expresses diminutiveness in this same, synthetic, way by means of suffixation. Recent works on English word-formation seem to deal somehow perfunctorily with diminution; thus, for example, Plag (2003: 13, 120-121) comments on diminution only in terms of pet forms where clipping is combined with suffixation. Adams (2001: 55-58) covers diminutives in more detail, listing eight nominal suffixes with (partly) diminutive meaning. Further, Bauer (2002: 1677-8) lists five regular suffixes, two prefixes, and a few “irregular” or historical forms. Schneider (2003: 78), on the other hand, enumerates as many as 86 different formatives that can be classified as diminutive suffixes of English, but, admittedly, many of these are only rarely used or are decidedly foreign in origin. Schneider therefore analyzes the following fourteen suffixes as present-day diminutive English suffixes (see Schneider 2003: 85 ff): *-a*, *-een*, *-er*, *-ette*, *-ie/-y/-ey/-ee*, *-kin*, *-le*, *-let*, *-ling*, *-o*, *-peg*, *-poo(h)*, *-pop*, *-s*. Like in Slovene, the suffix is attached to a nominal base to produce a diminutive form:

[small] kitchen > kitchen-ette	(kitchenDIM)
[baby] boot > boot-ee	(bootDIM)

In some other Slovene diminutive lexemes (e.g. *metuljček*, *ježek*, *polžek*, *miška*), the regular suffixation from the underlying paraphrase as described above (e.g. [majhen] metulj[-ø] – [small] butterfly) is further developed by means of terminological suffixation from a syntactic base following the pattern [tisti, ki je kot] metulj[-ø] ([that which is like a] butterfly), producing the term *metuljček* in the meaning of ‘bow-tie’. This can lead us to conclude that connotative similarity and productivity of the suffix which is used with the nominal base are crucial in the transition of a lexeme into a term.

From the point of view of lexicographic treatment in the *SSKJ*, diminutives – lexemes (i.e. dictionary entries in diminutive form) are even more interesting. The same morphological form can express denotative as well as connotative meaning (and use). In these cases the importance of the context is more clearly emphasized, e.g. *črvič* – *črviček* meaning ‘small child’, *gospodek* – *gospodič* meaning ‘young gentleman’, and this feature is also noticeable in English, as can be exemplified by the following illustrative contexts:

Dolgo je pazila na tega črvička. – She took care of that helpless little baby for a long time.  
Ta gospodič je zelo zahteven. – The young gentleman is most demanding.

Within diminutiveness in Slovene gradation of the diminutive quality can be expressed:

*fant* [boy] > *fant-ič* [boyDIM] > *fantič-ek* [boyDIMDIM]  
*hip* [moment] > *hip-ec* [momentDIM] > *hipč-ek* [momentDIMDIM]  
*kmet* [farmer] > *kmet-ič* [farmerDIM] > *kmetič-ek* [farmerDIMDIM]

Such multiple diminution occurs very rarely in English, and if it does, it is formed analytically and is marked in emotionality (e.g. *a tiny little cottage* as used in fairy tales).

### 3 TYPOLOGY OF DIMINUTIVES IN SLOVENE

Within word-formational stylistics in Slovene (Vidovič Muha 1986, 1988, 1995), nominal diminutives can be described from different viewpoints:<sup>6</sup>

1. the underlying paraphrase contains a unit with the denotative meaning ‘small, young’

mlad fant > *fant-ič/fant-ek*,  
nedorasel piščanec > *pišč-e*,  
majhen čevelj > *čevelj-č-ek*.

or a unit with the connotative meaning ‘emotionally positive’ or ‘emotionally negative’

ljubek/majhen stric > *strič-ek*,  
prijazna/dobra stvar > *stvar-ca*,  
majhna/ljubka grdoba > *grdob-ica*,  
slab/majhen človek > *človeč-ek*.

2. the markedness of the base is not necessarily transferred to the derived form, as in:

baraba (negative) > *barab-ica* (endearing),  
norec (negative) > *norč-ek* (endearing),  
drek (negative) > *drek-ec* (euphemistic),  
napaka (negative) > *napak-ica* (euphemistic).

3. within the category of diminutiveness, three Slovene suffixes are marked for either positive or negative value:

*-ek (griček)*:

Ta griček je res lep. – That little hill is indeed pretty.

A samo na ta griček si zlezel? – You have only climbed that small hill?

*-ica (knjižica)*:

To je koristna knjižica. – This is a useful little book.

V vsem življenju je napisal samo to knjižico. – In his entire life, he has only produced this piffling book.

*-e (fante, človeče, siromače, mače)*

4. Double diminutiveness or gradation of the diminutive quality is expressed by means of double suffixation; such gradation intensifies the markedness or expressiveness of the diminutive:

---

<sup>6</sup> The examples cited here have been taken from the digitalized corpus at the Institute for Slovene Language SRC SASA (Gložančev et al. 2009); <http://bos.zrc-sazu.si> (accessed June 2011).

*fant-ič-ek,*  
*otroč-ič-ek,*  
*perut-nič-ica*

5. In most cases the grammatical features (gender and number) of the base are transferred to the diminutive. Exceptions to this rule can be found in the following cases:

žrebe (n) > *žreb-ič(-ek)* (m); tele (n) > *tel-ič(-ek)* (m)  
vino (n) > *vince* (n) > *vinček* (m); mleko (n) > *mlekce* (n) > *mlekec* (m)  
dekle (n) > *deklica* (f) > *deklič* (m) / *dekliček* (m); punca (f) > *punčka* (f) > *punček* (m)  
reva (f) > *revše* (n); krava (f) > *kravše* (n)

As for the English language, the distinctions of the above kind are not to be found in English synthetic diminutives, but need to be expressed differently and are often found expressed elsewhere in a specific, actual context, as can be inferred from the sample sentences under 3. above.

#### 4 DIMINUTIVES AS LANGUAGE METAPHORS OR TERMS

The word-formational and semantic productivity of the Slovene diminutive suffixes such as *-ek* and *-ica* and the existence of ‘productive’ nouns based on metaphoric (associative) extension are a good starting point for the formation of terminology. In such cases, the diminutive as language metaphor gains semantic independence and becomes a technical term.

Regular suffixation, based on the pattern [tist-i/-a/-o kot] (‘the one like’) with its encoded analogy creates conditions for the emergence of the so-called language metaphor, which is also a semantically independent term, e.g.

*besedica* (linguistic)<sup>7</sup> [grammatical word],  
*blazinica* (zoological) [pad],  
*bobnič/bobenček* (medical) [eardrum],  
*cekinček* (zoological) [Scarce copper, *Heodes virgaureae*],  
*dimek* (botanical) [hawksbeard, *Crepis*],  
the two forms *dimec* and *dimček*, however, remain unmarked terminologically.

It is interesting that sometimes diminutive forms of both degrees are used terminologically, e.g. *čebrič/čebriček* (gastronomical), *dežnik/dežniček* (botanical), *kolence/kolenček* (botanical). Grammatical differences, however, may sometimes lead to terminological differentiation, as in *copatek* (m) (ballet shoe) as opposed to the non-terminological expression *copatka* (f) (small slipper).

---

<sup>7</sup> The labels in brackets for terminological fields are quoted after the *Slovar slovenskega knjižnega jezika* (SSKJ) I-V.

When equivalent English terms for Slovene lexicalized diminutives are sought for, it is very rare that the corresponding English terms also have diminutive suffixes, as in the pair *bradica* – *goatee* or *kipec* – *statuette*. The equivalents in English are in most cases lexemes with no diminution expressed in them, as is clearly shown by the examples quoted above.

## 5 CONCLUSION

An analysis of the diminutives confirms that syntax and word-formation can complement each other transformationally via the meaning, and this can be observed inter- as well as intralinguistically. Apart from that, diminutives also point out to the potential of semantic expansion of the already existing lexemes.

### Bibliography

- ADAMS, Valerie (2001) *Complex Words in English*. Harlow, Essex: Pearson Education Ltd.
- BAJEC, Anton (1950–1959) *Besedotvorje slovenskega jezika I–IV*. Ljubljana: SAZU.
- BAUER, Laurie/Rodney HUDDLESTON (2002) “Lexical word-formation.” In: G. K. Pullum/ R. Huddleston (eds), 1621–1721.
- GAJDA, Stanisław (ed) (1995) *Języki słowiańskie 1945–1995. Gramatyka – leksyka – odmiany*. Opole: Uniwersytet opolski, Instytut filologii polskiej.
- GLOŽANČEV, Alenka/Primož JAKOPIN/Mija MICHELIZZA/Lučka URŠIČ/Andreja ŽELE (eds) (2009) *Novejša slovenska leksika: (v povezavi s spletnimi jezikovnimi viri)*. Ljubljana: Založba ZRC, ZRC SAZU.
- KLINAR, Stanko (1996) “Prevajanje slovenskih manjšalnic.” In: S. Klinar (ed), 194–212.
- KLINAR, Stanko (ed) (1996) *Prispevki k tehniki prevajanja iz slovenščine v angleščino*. Radovljica: Didakta.
- PLAG, Ingo (2003) *Word-Formation in English*. Cambridge: CUP.
- PULLUM, Geoffrey K./Rodney HUDDLESTON (eds) (2002) *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- SCHNEIDER, Klaus P. (2003) *Diminutives in English*. Tübingen: Max Niemeyer. *Slovar slovenskega knjižnega jezika (SSKJ) I–V*, 1970, 1975, 1979, 1985, 1991. Ljubljana: DZS.
- STRAMLJIČ BREZNIK, Irena (2007) *Besedotvorna kategorija samostalniških manjšalnic v Pleteršnikovem slovarju in v SSKJ*. Zora 57. Maribor: Filozofska fakulteta UM.
- STRAMLJIČ BREZNIK, Irena (2010) *Tvorjenke slovenskega jezika med slovarjem in besedilom*. Zora 71. Maribor: Filozofska fakulteta UM.
- TOPORIŠIČ, Jože (1976) *Slovenska slovnica*. Maribor: Obzorja.
- VIDOVIČ MUHA, Ada (1971/72) “Oris dveh osnovnih pojavov sistema knjižnega jezika.” *Jezik in slovstvo* 17/6, 178–186.
- VIDOVIČ MUHA, Ada (1972) *Kategorizacija in stilna opredelitev ozko knjižne leksike*. VIII. SSJLK. Ljubljana: Filozofska fakulteta, 35–52.
- VIDOVIČ MUHA, Ada (1986) *Besedotvorni pomen in njegova stilistika*. XXII. SSJLK. Ljubljana: Filozofska fakulteta, 79–91.

- VIDOVIČ MUHA, Ada (1988) *Slovensko skladijsko besedotvorje ob primerih zložen.* Ljubljana: Znanstveni inštitut Filozofske fakultete and Partizanska knjiga.
- VIDOVIČ MUHA, Ada (1995) "Dva tipološka zgleđa normativne vrednosti slovenske besedotvorne morfematike." In: S. Gajda (ed), 153-165.
- VIDOVIČ MUHA, Ada (2000) *Slovensko leksikalno pomenoslovje. Govorica slovarja.* Ljubljana: Znanstveni inštitut Filozofske fakultete. *Zbirke Inštituta za slovenski jezik Frana Ramovša ZRC SAZU.* June 2011. <http://bos.zrc-sazu.si>
- ZIMA, Jaroslav (1961) *Ekspresivita slova v současně češtině.* Prague: Filozofická fakultá.

#### Abstract

### NOMINAL DIMINUTIVES IN SLOVENE AND ENGLISH

This contrastive treatment of diminutives in Slovene and English aims to comment on the productivity of diminutive formations in both languages; in Slovene diminutives are particularly productive also in the field of terminology. Diminutiveness can be expressed in both languages in different ways: in Slovene lexemic word-formation is prevalent, while in English the same content is usually expressed syntactically. However, a combination of both is possible as well, as for example in *hišica – majhna hiša – majhna hiška*. A dual nature of diminutiveness can also be detected in graded cases like *hiša – hiška – hiškica, hči – hčerkica*. In English the formation of diminutives is less productive, however, the expression of diminutiveness is thus transferred to the syntactic level and is, as such, limitless. Since Slovene is a highly morphemic language it is only natural that diminutiveness should be expressed morphemically. The aim of the article is to highlight some differences in the ways diminutiveness can be expressed in both languages by means of examples.

#### Povzetek

### SAMOSTALNIŠKE MANJŠALNICE V SLOVENŠČINI IN ANGLEŠČINI

Kontrastivna obravnava manjšalnic v slovenščini in angleščini izkazuje porast tovrstnih tvorjenk, zlasti v slovenščini so zastopane tudi v terminologiji. Sicer se manjšalnost v obeh jezikih izraža na več različnih načinov; v slovenščini prevladuje leksemsko izražanje, v angleščini pa skladijsko, možna je tudi kombinacija obeh, npr. *hišica – majhna hiša – majhna hiška*. Dvojnost izražanja manjšalnosti se kaže v primerih kot *hiša – hiška – hiškica, hči – hčerkica – hčerkica*. V angleščini je sama tvorba manjšalnic precej manj produktivna, vendar je zato izražanje manjšalnosti preneseno na skladijsko raven in tako neomejeno. Ker je slovenščina izrazito morfemski jezik, je tudi manjšalnost izražena morfemsko. Namen razprave je s primeri poudariti ravno razlike izražanja manjšalnosti v obeh obravnavanih jezikih.

## LA MORPHOLOGIE MARGINALE DANS LE SLAM : COMPARAISON INTERLINGUALE FRANCE VS ALLEMAGNE

### 1 INTRODUCTION

La morphologie extra-grammaticale ou marginale en langues française et allemande mérite tout particulièrement que nous nous y attardions plus précisément. Il est évident qu'il s'agit là d'un vaste sujet qu'il nous faut restreindre à un type de corpus identique dans les deux langues afin d'obtenir une comparaison interlinguale de type quantitative. Nous souhaitons commencer cette étude par la présentation de ce que peut être le « Poetry Slam » ainsi que son expression dans les deux pays nous préoccupant ici. Nous continuerons cet article par la présentation de notre corpus pour ensuite passer dans le vif du sujet, à savoir la morphologie marginale en langues française et allemande et les résultats concrets liés à l'analyse des corpus. Nous concluons sur les possibilités existantes de comparabilité des procédés marginaux selon les deux langues analysées.

#### 1.1 L'essor du Slam

Nous allons présenter ici quelques dates clés historiques mettant en lumière le développement du mouvement du Slam. Il faut préciser que dans le courant des années 1970, les mouvements poétiques sont à bout de souffle. Le « Poetry Slam » a permis de relancer les lectures poétiques en un mouvement fascinant qui n'aurait pas été envisageable par le passé.

Dès 1975, les poètes et metteurs en scène Miguel Algarin et Miguel Pinero fondent dans le quartier de New-York « East Side » le « Nuyoric Poets Cafe » dans lequel se produisent au début plutôt des poètes portoricains. En parallèle ont lieu des concerts de Jazz et des projets de théâtre. A cette suite en 1985, Marc Smith « slampapi » organise dans le quartier du « Bucktown »<sup>1</sup> à Chicago des performances poétiques avec musique de jazz en arrière plan. (Bylanzky 2002 : 159) Il met en place le « Chicago Poetry Ensemble » qui servira plus tard à tous de prototype. En 1986, il organise un show de poésie dans le club de jazz « Green Mill » composé de 3 parties :

---

\* *Adresse de l'auteur* : Institut für Angewandte Linguistik und Translatologie, Beethovenstraße 15, 04107 Leipzig, Deutschland. Mél : mberron@hotmail.fr

\*\* *Adresse de l'auteur* : Institut für Angewandte Linguistik und Translatologie, Beethovenstraße 15, 04107 Leipzig, Deutschland. Mél : mouton@rz.uni-leipzig.de

<sup>1</sup> Dans un café portant le nom de « Get Me High Lounge ».

- 1) Open mike,
- 2) Présentation des invités venus pour slamer de tous les USA,
- 3) Spectacle du « Chicago Poetry Ensemble ».

Ce concept étant trop contraignant dans la préparation, Marc Smith finira par instaurer une semaine sur deux une sorte de duel poétique. Il s'agit en réalité du premier Poetry slam<sup>2</sup> dans le monde. (Bylanzky 2002 : 159) Dès 1991/1992, à la suite du « National Poetry Slam » de Boston, on verra non seulement apparaître des clips vidéo de Poetry Slam à la télévision américaine mais la chaîne de télévision MTV proposera l'émission « Poetry unplugged ». La médiatisation sur les télévisions américaines a eu comme répercussion le développement des premiers Slam en Finlande et en Suède en 1993 ainsi qu'en Angleterre (Londres) sous la direction de John Paul O'Neil. Le « National Poetry Slam » en 1993 à Chicago a attiré des équipes en provenance du Canada et de la Finlande. Il est fascinant de constater l'impact du Slam sur le Japon qui organise ses premiers Slam dès 1994 à Tokyo. L'Allemagne fera de même un an plus tôt en 1993 à l'initiative d'un des pionniers du Slam « Wolfgang Hoge kamp » et de deux américains vivant sur le territoire allemand : Priscilla Be et Rik Maverik. En 1998 à Munich, le Slam est découvert par les adeptes du Hip Hop en effectuant la fusion du Hip Hop et de la littérature. Un phénomène d'internationalisation prend ensuite forme, en partie grâce au film de Marc Lévin « Slam »<sup>3</sup> faisant ainsi connaître ce phénomène de littérature orale à l'échelle internationale. Qui plus est, en 2002, nombreux sont les slameurs américains célèbres venant se produire sur les scènes allemandes tels que, entre autre, Marc Smith, Tantra, Aqiyl, Ainsley, Will Power, Soul Evans (Bylanzky 2002 : 170). On verra effectivement par la suite l'impact des anglicismes sur le Slam allemand et ceci pourrait être effectivement une explication intéressante même si d'autres facteurs rentrent également en compte. Le mouvement du Slam s'est implanté bien plus tard sur le territoire français même si l'on peut déjà remarquer dès 1995 un noyau dur composé de « Nada, Pilote le Hot, Joël Barazer » mêlant<sup>4</sup> ainsi des poètes, des performeurs et des rappeurs.<sup>5</sup> Le premier spectacle<sup>6</sup> de Slam en France fut joué aux rencontres urbaines de la ville en novembre 2001 à Paris. Le film de Marc Lévin cité précédemment a accéléré le développement du Slam en France à partir de 1998. Pilote le Hot<sup>7</sup> fut l'un des tout premiers à mettre sur pied des scènes de Slam dans l'Est parisien. Courant 2000, le mouvement se développe et s'implante tout particulièrement à Paris et dans sa banlieue (129H, Spoke orchestra), avant de se répandre dès 2002

---

<sup>2</sup> Marc Smith le nommera « The uptown Poetry Slam ».

<sup>3</sup> Le film a été tourné en 1997 avec le Slameur Saül Williams comme acteur principal.

<sup>4</sup> Dans un bar de Pigalle à Paris « Le Club club ».

<sup>5</sup> [www.planeteslam.com](http://www.planeteslam.com)

<sup>6</sup> Intitulé « Les gens et moi » mis en scène par Gérard Mendy et dans lequel on retrouve le slameur Nebil Daghzen, né à Paris en 1980.

<sup>7</sup> Né le 10/04/1966 (Paris - Aubervilliers).



dans les grandes villes de province telles que Reims, Lille, Nantes, Rennes, Tours... En 2004, de grandes rencontres sont organisées un peu partout en France: Nuit du Slam (Reims, Creil, Lyon, Dijon), Slam Fever (Rennes), Slam L'homme géant (Lyon), Slam N'co (Nantes), Super Slam (Tours) et le Grand Slam de Paname (Paris).

Ainsi, nous pouvons constater en conclusion que Marc Smith, passionné de lecture poétique, a ainsi cherché à déclencher un phénomène de lecture poétique à l'égal des manifestations de Jazz. Son but était également d'abolir les frontières entre les poètes de la rue et les poètes académiques, de réussir à mélanger les styles et genres poétiques. Le Slam est donc considéré (cf. Preckwitz 2002 : 43) comme la forme la plus développée du « Spoken Word ». Ainsi, le Slam est une littérature trouvant son origine au sein de la population, étant pensé et créé pour la population.

## 1.2 Définition du Slam

Du point de vue de la traduction directe, le terme du « Slam » demande une explication plus appropriée. Le terme « Slam » a de nombreuses définitions différentes dans les domaines de la langue anglaise et américaine. Il est important de préciser que le mot « Slam » provient de l'ancien Scandinavien (XVIII<sup>e</sup> siècle) « Slamra, slambra » (Preckwitz 2002: 20) et cela avant de prendre la signification de « frapper, claquer » après traduction directe de l'anglais. De nos jours dans le domaine sportif, le tennis par exemple, on utilise ce mot pour désigner un tournoi; dans les jeux de carte, il signifie un pli et en Amérique le slameur est un prisonnier. Depuis 1984, il représente un tournoi poétique compétitif sous une forme particulière (cf. Anders 2004 : 14). Important à différencier sont également les nuances de signification entre:

- 1) « poetry slam » en tant que performance compétitive,
- 2) « le Slam » en tant que mouvement littéraire,
- 3) « slam poetry » en tant que forme de compétition orale tournée vers le public.

Le Slam affirme le verbe comme force rédemptrice et montre une nouvelle voie pour le Hip Hop, prenant ses distances avec le gangster Rap. Le Slam possède de nombreux points communs avec le domaine musical du Rap et du Jazz. Tout d'abord, les slameurs utilisent des éléments relatifs au Jazz comme la strophe, le refrain, l'improvisation, réponses en simultanée ou non envers d'autres slameurs. Les slameurs reprennent également les grandes thématiques du Rap comme les thèmes sociaux critiques possédant un rythme caractéristique parfois agressif. On voit également beaucoup d'artistes de Rap prendre part aux scènes de Slam. En général, les rappeurs font souvent partie d'un groupe social ethnique et la musique est pour eux une projection de la société, tandis que le Slam regroupe diverses formes ainsi que des slameurs d'origines aussi diverses que variées. Le Rap est donc un type de musique à texte, et plus précisément le retour au texte dans la tradition de la chanson française, à la suite, par exemple, de la musique de Georges Brassens (cf. Dannhorn 2002 : 70).

### 1.3 L'expression du Slam en France et en Allemagne

La structure du système américain repose essentiellement sur le modèle compétitif avec l'attribution d'un jugement après la performance du Slameur. Les scènes de Slam y sont très ritualisées et composées d'un organisateur, d'un maître de cérémonie, d'un jury tiré au sort parmi le public présent dans la salle, les auteurs et le reste du public. Le jury attribue après chaque séquence une note avec justification de leur décision. Contrairement aux scènes de Slam américain, les Français ont exclu le jury du déroulement de la soirée pour ne garder que le concept suivant:

*« On est slameur cinq minutes et après on est spectateur »  
« Un slam, un verre »<sup>8</sup>*

En règle générale, les scènes de Slam en France sont considérées comme ouvertes, même si l'on retrouve également le Slam sous forme de tournoi. Le concept de scènes ouvertes signifie que chacun peut décider de venir sur scène afin d'y déclamer un texte<sup>9</sup>. On y observe donc un Maître de cérémonie (MC) qui joue le rôle de modérateur. Il ne lit pas lui-même de textes même si cela peut lui arriver lors d'un manque de participation de la part des Slameurs. Il n'y a ni gagnant ni perdant. Chaque participant reçoit une boisson gratuite, dans la mesure où la scène se passe dans un café ou dans un bar. L'entrée est le plus souvent gratuite bien qu'il y ait également des scènes payantes. Les Slameurs ne payent pas leur entrée. Les organisateurs du Slam en Allemagne ont fortement repris le modèle américain qui considère ce type de lecture poétique comme une compétition poétique. Même si le modèle américain est encore plus fortement axé sur le modèle compétitif – en effet des sommes d'argent sont fréquemment offertes aux vainqueurs du tournoi – il y a un maître de cérémonie, qui joue le rôle de modérateur et qui va également choisir le jury parmi le public. Le jury peut être composé du public dans son ensemble ou bien d'une fraction représentative de 5 à 10 personnes. Le jugement se fait par vote sous les différentes formes suivantes :

- 1) Applaudissement,
- 2) main levée,
- 3) note allant de 0 à 10,
- 4) bulletins secrets.

Les participants passent par groupe de cinq, on assiste au vote du meilleur des cinq participants, puis on passe aux cinq participants suivants. En général, il y a 3 à 4 roulements de cinq participants, puis les gagnants repassent pour désigner le vainqueur final. Il y a un prix pour le gagnant : il peut s'agir d'une bouteille d'alcool

---

<sup>8</sup> Site Ubackconcept.

<sup>9</sup> Ce qui était à la base le concept « Open Mike » de Marc Smith.

mais la plupart du temps, on offre un ouvrage littéraire. Les règles du tournoi de poésies instaurées par Marc Smith sont les suivantes<sup>10</sup>:

- 1) Micro ouvert à tous sans aucune forme de discrimination.
- 2) Trois minutes de temps de parole.
- 3) Un poème par passage sans musique, ni costume, ni accessoire.
- 4) Un jury populaire de cinq personnes pris au hasard dans le public pour noter les performances.
- 5) Une finale offrant un deuxième passage aux poètes totalisant les plus forts scores.

Il semble évident que des différences se sont introduites dans le concept de base même si la base reste la même. De même selon l'endroit, selon la scène et selon les organisateurs, des règles vont être modifiées et adaptées – il n'y a donc pas de définition figée de la forme prise par le Slam en France ou en Allemagne. Une autre différence fondamentale présente entre l'Allemagne et la France : les Slameurs allemands (du corpus) ont rarement recours à un fond musical en arrière plan de leur texte ou même à un style tournant vers le Rap ou le Hip Hop tandis que les Slameurs français (du corpus) se situent parfois à la frontière des genres. Ainsi, on peut donc malgré tout en conclure que la forme américaine du Slam s'est non seulement imposée plus tôt en Allemagne (1993) qu'en France (1998-2000) mais également plus fortement dans sa tradition.

## 2 LES PROCÉDÉS FORMELS EN MORPHOLOGIE MARGINALE

Les différents procédés formels que nous voulons présenter dans le cadre de cette analyse pratique peuvent être catégorisés selon un grand nombre d'étiquettes différentes telles que les notions de morphologie extra-, non ou agrammaticale, marginale, mineure, périphérique, créative, expressive, non conventionnelle (Fradin 2007 : 1). Il s'agit donc dans ce type de procédés de phénomènes linguistiques non conformes au système linguistique. Il est très fréquent d'après Fradin (2007 : 1f) d'utiliser la notion d'extra-grammaticalité ou de marginalité pour parler de ce qui n'est pas régulier, de ce qui sort de l'ordinaire. La langue n'est pas seulement composée de lexèmes présents dans le dictionnaire et l'on retrouve bon nombre d'emprunts à des langues étrangères ou à des sociolectes, langues de spécialité et des créations occasionnelles. Ce sont ces lexèmes qui sont tout particulièrement sujets à des écarts formels (Fradin 2007 : 4). Comme l'indiquent très justement Kilani-Schoch et Dressler (2005 : 98):

*« La morphologie extra-grammaticale (...) désigne un ensemble de phénomènes hétérogènes qui ont pour unique caractéristique commune de ne pas respecter l'un ou l'autre principe de la morphologie grammaticale. »*

---

<sup>10</sup> [www.ligueslamdefrance.com](http://www.ligueslamdefrance.com).

Fradin (2007 : 4f) donne dans son article quatre critères permettant de classer une unité lexicale comme étant extra-grammaticale :

- 1) Le caractère conscient de la création comme dans le cas des langages secrets même si le procédé en lui-même n'est pas forcément compris.
- 2) L'extension typologique du procédé.
- 3) La productivité du procédé comprenant la production de lexèmes qui n'ont jamais existé auparavant.
- 4) L'interaction existante entre les différents modules de grammaire.

Les différents procédés de la morphologie marginale auxquels Fradin (2007) fait allusion sont entre autres : les hypocoristiques, les tronqués suffixés en *-o* tels que *intello*, *hystéro* (Kilani-Schoch et Dressler 1992), les acronymes ou sigles, réduction (apocope, aphérèse), les mots-valises, les réductions, les langages secrets tels le *verlan*. Comme nous le constatons dans cette liste qui ne se veut pas exhaustive, les procédés ayant trait à la morphologie extra-grammaticale voire marginale sont en grande partie les procédés dits formels. Les procédés que nous présenterons ici sont tous basés sur les caractéristiques utilisées dans la langue populaire et argotique française. Nous y rajoutons également les procédés typiques de l'usage « jeune » des cités (FCC), de même en langue allemande les jeunes ont essentiellement recours au modèle de formation de la grammaire allemande, on y observera malgré tout des aspects typiquement jeunes.

Nous tenons à préciser ici que dans le cadre de notre comparaison interlinguale, nous choisissons comme langue de départ la langue française qui sera ainsi la base de notre comparaison.

Comme dit précédemment, les procédés utilisés dans la formation des mots des argots sont ceux qui sont utilisés dans la grammaire standard du français et de l'allemand mais en plus grande quantité et avec une volonté créative de former de nouveaux mots. Quelques particularités sont à observer dans le tableau ci-dessous :

Procédés formels français	Procédés formels allemands	Procédés formels en commun
La dérivation Préfixe Suffixe Suffixation parasitaire	La dérivation Préfixe Suffixe (-italien) Faux suffixe Dérivation parasynthétique Dérivation inverse	La dérivation
Troncation Par apocope Par aphérèse	Troncation Par apocope	Troncation
Resuffixation après apocope		
Redoublement hypocoristique	Redoublement hypocoristique	Redoublement hypocoristique
Composition Verbe + nom (le plus souvent)	Composition Nom + nom (en particulier avec des anglicismes) Avec affixoïde	Composition
Abréviation	Abréviation	Abréviation
Siglaïson	Siglaïson	Siglaïson
Agglutination/déglutination		
Calembour Attraction paronymique Mots-valises	Contamination Mots-valises	Calembour Mots-valises
Onomatopée	Onomatopée	Onomatopée
Verlan		
Largonji		

Ce tableau nous indique clairement les différences majeures apparaissant entre les procédés formels allemands et français : l'absence du verlan est à remarquer, l'une des catégories majeures de l'argot des jeunes en France. Mais comme l'indique Fradin dans son article de 2007, le verlan de par son système d'inversion ne se retrouve qu'en langue française. On remarquera ici son absence flagrante en langue allemande et dans le corpus allemand actuel enregistré sur le terrain et sa présence marquante dans le corpus français. Le verlan est l'un des processus de l'argot des jeunes ou du FCC, mais a également sa place dans l'argot commun selon son degré d'intégration dans la population. Il est évident que des différences apparaissent également dans le cadre des dérivations ou des troncations mais le principe de base reste malgré tout le même.

### 3 ANALYSE LINGUISTIQUE DU CORPUS

La première partie de notre recherche de corpus s'est déroulée en France dans la ville de Lille entre septembre 2008 et avril 2009, tandis que la deuxième partie de notre recherche de corpus s'est déroulée en Allemagne dans la ville de Berlin entre

septembre 2009 et septembre 2010. Ce corpus conséquent comprend 218 textes allant de 2 à 7 minutes selon les textes. Le corpus français comprend un total de 137 textes analysés tandis que le corpus allemand comprend 81 textes dont, actuellement, 37 ont déjà été analysés. Nous nous appuyons ci-dessous sur les textes du corpus analysés pour tirer des conclusions concernant le sujet de notre article.

### 3.1 Exemples tirés du corpus français

Nous présenterons ici les exemples les plus probants de la morphologie marginale et tout particulièrement des procédés formels tels que le verlan, les troncations par aphérèse ou par apocope, la suffixation ou resuffixation après apocope, les sigles, la réduplication pour ne citer que ceux se retrouvant dans le corpus analysé.

Nous commencerons par présenter une partie des exemples de verlan (langage secret) présents dans le corpus : Nombreux sont les exemples de verlan de type simple S1S2 – S2S1 tels que *la teuté* - tête, *veugra* - grave, *la ne-doua* - la douane, *ive* - vie, *les renpas* - les parents, *asskom* - comme ça, *le gen-ar* - l'argent, *le ceucra* - le cracre, *le chromi* - le micro, *québlo* - bloqué, *un skeudi* - un disque, *teubé* - bête, *le rollbas* - le bas rôle. Certains exemples de verlan du corpus sont également suivis d'une préfixation comme dans *la areuga* - la gare avec une préfixation en *-a*, ou encore dans *la incarna* - l'arnaque avec une préfixation en *-in*. Ces deux exemples sont intéressants et basés non seulement sur le principe du verlan mais avec un ajout créatif de la part des auteurs de façon à colorer l'expression utilisée. Des troncations par apocope sont aussi assez fréquentes comme dans le cas de *Skeud* - *disque* - *skeudi*, ou de *DAS* - *sida* - *DASI*. Comme nous le verrons dans la suite de cet article, le corpus français contient bon nombre de troncations par apocope tels que *la connexe* - la connexion, *une instru* - une instrumentale, *le mic* - le micro, *la resign* - la résignation, *un indic* - un indicateur, *une impro* - une improvisation. Les sigles tels que *être OP* - être open, *le MC* - le maître de cérémonie, apparaissent également mais en faible quantité. Certaines réductions de type onomatopée sont intéressantes comme *Beng beng* qui vient de *beng beng* symbolisant le son rendu par le djumbé africain suivi d'une resuffixation en *-a* sans doute pour reproduire d'une manière plus musicale et pour accentuer le son rendu par le tambour.

### 3.2 Exemples tirés du corpus allemand

Le corpus allemand présente des phénomènes courants de troncation par apocope et par aphérèse, de suffixations basées sur des modèles courants de la grammaire allemande mais utilisées de façon marginale dans le corpus, d'affixoides typiques de la langue allemande des jeunes, de sigles, de réductions et d'onomatopées. Par exemple *das Klo* provient de la troncation par apocope de *Klosettpapier* ayant la signification de *papier toilette* ou encore l'exemple de *Öko* provenant de la troncation par apocope de *Ökologisch*. La suffixation est intéressante de par sa productivité. Ainsi, nous allons analyser ici quelques exemples marginaux présents dans le corpus: tout d'abord le terme de *Dreisam* ayant la signification d'*être solitaire* à

*trois* – ce terme est absent du dictionnaire. *Drei-* signifie *trois*, et le suffixe *-sam* est emprunté au terme de *einsam* ayant la signification de *être solitaire*. Les suffixes en *-er* ont en allemand la signification de *quelqu'un qui fait quelque chose*. Le terme de *Sparpreiser* a la signification de *quelqu'un qui fait des économies en achetant un billet de train peu cher* – le suffixe *-er* a été rajouté au substantif de base *Sparpreis* ayant la signification de *billet de train en promotion*. Ce type de suffixation montre un procédé de la langue allemande utilisé à un emplacement inconnu par les dictionnaires ou les utilisations habituelles. Les affixoïdes ont une place importante dans la langue des jeunes en Allemagne – c'est une source de création très productive comme dans les exemples suivants *Megafick*, *Monsterfick* ou encore *affenfick*<sup>11</sup> dont la signification se déduit d'elle-même. Les sigles sont en règle générale très simples tels que *LSD*<sup>12</sup> ou *DNT*<sup>13</sup> mais sont parfois associés à une resuffixation provoquant par là même un jeu de mot comme dans le cas de *DJANE*. Ce dernier associe le sigle *DJ* – *disque jockey* à la suffixation en *-ane* permettant ainsi à la création du prénom *Jane* avec une préfixation en *-d*. Le terme de *Mumu* provient du substantif *Muschi*<sup>14</sup> ayant suivi non seulement une troncation par apocope mais qui plus est une réduplication. Outre les onomatopées de type simple telles que *Hoho*, *Tick Tack*, *tsch tsch tsch* – des onomatopées suivies d'une composition afin de donner un aspect plus imagé de la situation décrite par le locuteur sont présentes dans le corpus : *Tick-Augen*<sup>15</sup> ou *Tick-Blick*<sup>16</sup>. Pour terminer, nous présenterons ici un mot-valise : *Extpotenziell* provenant de *Externe Potenziell* avec la signification de *potentiel externe*.

### 3.3 Comparaison des procédés formels dans les deux langues

Les deux corpus présentent dans la plupart des cas des phénomènes de type formel que nous pouvons considérer comme équivalents. Il est évident que le procédé formel du verlan est absent de la langue allemande, mais l'on retrouve des procédés communs de troncation par apocope ou par aphérèse, de sigles, de réductions, d'onomatopées et de suffixations. Les mots-valises sont dans la formation tout aussi présents en langue française qu'allemande même si notre corpus français n'en présente qu'un exemple et le corpus allemand n'en présente au stade de notre recherche que deux exemples concrets. Les affixoïdes<sup>17</sup> sont présents dans le corpus allemand mais pas dans le corpus français. Les tableaux suivants nous permettront

<sup>11</sup> *Affen* signifie singe.

<sup>12</sup> Acide lysergique diéthylamide (voire aussi SPEED).

<sup>13</sup> Dysembrioblastische neuroepitheliale Tumore.

<sup>14</sup> La moule – représentation du sexe féminin.

<sup>15</sup> *Augen* signifie les yeux.

<sup>16</sup> *Blick* signifie le regard.

<sup>17</sup> Appelés également « *halb-affix* » en langue allemande. Ils ne sont plus des affixes mais ne peuvent pas encore être employés seuls.

d'obtenir une meilleure vue d'ensemble du pourcentage des procédés marginaux utilisés par les slameurs de notre corpus selon la langue :

<b>Procédés</b>	<b>Nombre d'occurrences dans le corpus français</b>	<b>Nombre d'occurrences dans le corpus allemand</b>
Troncation par apocope	37 34,6%	4 4,3%
Troncation par aphérèse	4 3,7%	2 2,1%
Troncation + resuffixation ou Suffixation simple	17 (suffixes) 16,0%	18 (10 suffixes / 8 préfixes) 19,1%
Verlan simple	16 15,0%	0 0%
Verlan + troncation	7 6,5%	0 0%
Troncation ou suffixation + verlanisation	3 2,8%	0 0%
Affixoïde avec composition	0 0%	5 5,3%
Composition N+N	4 3,7%	28 29,8%
Composition A+N	1 0,9%	2 2,1%
Composition A+V/ N+V+N	0 0%	8 8,5%
Sigles	2 1,9%	2 2,1%
Mot-valise	1 0,9%	2 2,1%
Réduplication (redoublement hypocoristique)	8 7,5%	2 2,1%
Réduplication + suffixation	1 0,9%	0 0%
Onomatopée	6 5,6%	21 22,4%
<b>TOTAL</b>	<b>107</b> <b>100%<sup>18</sup></b>	<b>94</b> <b>100%</b>

<sup>18</sup> Les valeurs en pourcentage sont arrondies.



Les résultats chiffrés du tableau ci-dessus nous permettent de tirer les conclusions suivantes : il existe deux catégories non représentées dans les deux corpus, d'une part le verlan qui n'est présent que dans le corpus français (24,3 %) et d'autre part les affixoides (5,3 %) qui n'apparaissent que dans le corpus allemand. À observer également l'importance prise par le procédé des troncations par apocope dans le corpus français avec 34,6 % des occurrences pour seulement 4,3 % dans le corpus allemand. Les compositions sont également fortement représentées dans le corpus allemand avec 40,4 % des occurrences pour 4,6 % des occurrences dans le corpus français. Les auteurs du corpus français axent plus fortement sur les redoublements hypocoristiques (7,5 %) que les auteurs allemands (2,1 %) mais ces derniers ont plus recouru aux onomatopées (22,4 %) que les auteurs du corpus en langue française (5,6 %).

#### 4 CONCLUSION GÉNÉRALE

Pour conclure, le Slam est une littérature orale populaire ouverte et accessible à tous. Il s'agit avant tout d'une poésie sortant des sentiers battus, d'où l'importance de la créativité lexicale en présence. Il est évident que les phénomènes de la morphologie marginale peuvent être considérés comme équivalents même si l'on repère des différences fondamentales comme le verlan en langue française ou encore les affixoides en langue allemande. Les autres phénomènes formels sont, dans le corpus de notre analyse, présents mais dans des proportions différentes. Ces différences de proportion sont explicables par leur fréquence dans chacun des deux langues courantes respectives.

#### Bibliographie

- ANDERS, Petra (2004) *Poetry-Slam, live-poeten in Dichterschlachten- Ein Arbeitsbuch*. Mühlheim an der Ruhr: Verlag an der Ruhr.
- ANDROUTSOPOULOS, Jannis K. (1998) « Forschungsperspektiven und Jugendsprache: Ein integrativer Überblick. » In : J. K. Androutsopoulos/A. Scholz (Éds.): *Jugendsprache, langue des jeunes, youth language*. Francfort/Main: Peter Lang. Europäischer Verlag der Wissenschaften, 1-34.
- BERRON, Marie-Anne (ouvrage encore non publié) *L'Argot dans le SLAM – Une analyse linguistique comparée entre la France et l'Allemagne*. Thèse de doctorat : Leipzig.
- BARTHES, Roland (1953) *Le degré zéro de l'écriture*. Paris: Édition du Seuil.
- BYLANZKY, Ko (2002) *Planet Slam. Das Universum Poetry Slam*. Riemerling : Yedermann.
- BOYER, Henri (1997) « 'Nouveau Français', 'Parler jeune' ou 'Langue des cités' ? Remarques sur un objet linguistique médiatiquement identifié. » *Langue française* 114, 6-15.
- BOURDIEU, Pierre (2004) *Ce que parler veut dire*. Paris: Fayard.
- BOURDIEU, Pierre (1991) « Vous avez dit "populaire" . » In : *Actes de la Recherche en Sciences Sociales*, Paris: Minuit, N° 90, 98-105.
- DANNHORN, Susanne/Adelheid Hu/Pauline LUQUIN-ELHILALI (2002) « Poésie d'amour française et musique pluriculturelle. » *Französisch heute* 1, 66-75.
- ERNST, Thomas (2005) *Popliteratur*. Hamburg: Eva/Europäische Verlagsanstalt.

- ESNAULT, Gaston (1965) *Dictionnaire historique des argots français*. Paris: Larousse.
- FRADIN, Bernard/Fabio MONTERMINI/Marc PLÉNAT (2007) « *Morphologie grammaticale et extragrammaticale*. » In : B. Fradin/F. Kerleroux/M. Plénat (éds) *Aperçus de morphologie du français*. Paris : Presses de l'Université de Vincennes, 21-45.
- FRANCOIS-GEIGER, Denise/Jean-Pierre GOUDAILLIER (1991) « Parlure argotique. » *Langue française* 90, 3-113.
- GADET, Françoise (2003) *La variation sociale en français*. Paris: Éditions Ophrys.
- GOUDAILLIER, Jean-Pierre (2002) « De l'argot traditionnel au français contemporain des cités. » *La Linguistique* 38, 5-23.
- GOUDAILLIER, Jean-Pierre (2001) *Comment tu tchatches! Dictionnaire du français contemporain des cités*. Paris: Maisonneuve et Larose.
- KERENSKY, Boris und STEFANESCU, Sergiu (Hrsg.) (1998) *Es gibt Social Beat/ Slam Poetry*, Stuttgart: Keltensberg.
- KILANI-SCHOCH Marianne/Wolfgang U. DRESSLER (1992) « Prol-o, intell-o, gauch-o et les autres. Propriétés formelles de deux opérations du français parlé. » *Romanistisches Jahrbuch* 43, 65-86.
- KILANI-SCHOCH Marianne/Wolfgang U. DRESSLER (2005) *Morphologie naturelle et flexion du verbe français*. Tübingen: Gunter Narr Verlag.
- NEUMEISTER, Andreas und HARTGES, Marcel (Hrsg.) (1996) *Poetry! Slam! Texte der Pop-Fraktion*, Hamburg: Rowohlt Taschenbuch Verlag GmbH.
- PRECKWITZ, Boritz (2005) *Spoken word und poetry slam*. Wien: Passagen-Verl.
- PRECKWITZ, Boritz (2002) *Slam poetry - Nachhut der Moderne*. Berlin: B. N. Breckwitz.
- SCHERFER, Peter (2003) « Jugendsprache in Frankreich. » In : E. Neuland (éd) *Jugendsprache - Jugendliteratur - Jugendkultur. Interdisziplinäre Beiträge zu sprachkulturellen Ausdrucksformen Jugendlicher*. Francfort/Main: Peter Lang, 141-168.
- SHÄFER, Jürgen (1997) *Pop-Literatur: Rolf Dieter Brinkmann und das Verhältnis zur Populärkultur in der Literatur der sechziger Jahre*. Siegen: Universität Dissertation.
- WESTERMAYR, Stefanie (2004) *Poetry Slam in Deutschland*, Marburg: Tectum Verlag.
- Le dictionnaire de la zone*. 2011. <http://www.dictionnairedelazone.fr/>.

## Résumé

### LA MORPHOLOGIE MARGINALE DANS LE SLAM: COMPARAISON INTERLINGUALE FRANCE VS ALLEMAGNE

Nous proposons à partir d'un phénomène de poésie contemporaine (Slam) d'analyser l'importance de la morphologie marginale (Fradin 2007) comprenant des phénomènes tels que, entre autres, le verlan, les troncations par apocope ou par aphérèse, les suffixations, la reduplication, les onomatopées, les sigles et les mots-valises. L'intérêt de cet article repose sur le fait de présenter ces différents éléments tirés d'un corpus oral rassemblé en France au cours de l'année 2008/2009 par le biais d'enregistrements de terrain, ainsi que de la volonté de proposer une comparaison interlinguale du même phénomène de poésie en langue allemande d'après un corpus rassemblé lui en 2009/2010. Cet article propose non seulement un repérage des procédés formels communs et ceux divergents de la morphologie marginale dans les deux langues, mais encore un pourcentage des différents procédés dans le corpus analysé.

## Povzetek

### MARGINALNA MORFOLOGIJA V SLAMU: FRANCOSKO – NEMŠKA MEDJEZIKOVNA PRIMERJAVA

Na primeru slama kot sodobnega poetskega izraza raziskujemo pomen marginalne morfologije (Fradin 2007), ki med drugim zaobjema pojave spreobračanja zlogov (*verlan*), krajsanja z apokopo ali aferezo, priponsko izpeljevanje, podvojitve, kratičenje in spajanje. Namen članka je bil v prvi fazi analizirati različne podatke, pridobljene iz govornega korpusa, ki sva ga s pomočjo terenskih snemanj zbrali v Franciji v letih 2008-2009, v drugi fazi pa te podatke primerjati z analizo podatkov, pridobljenih iz nemškega korpusa, ki sva ga zbrali v letih 2009-2010. Poleg določitve posameznih formalnih postopkov marginalne morfologije, ki so oziroma niso skupni obema jezikoma, članek ponuja statistično analizo v korpusu zastopanih morfoloških procesov.



## MEASURING MORPHOLOGICAL CANONICITY

### 1. INTRODUCTION

The question of regularity within morphological paradigms has been formerly addressed within approaches falling in the scope of *Canonical Typology*.

This work follows in the tradition of those approaches. It adopts a *strictly lexicalist approach* to morphology (Karttunen 1989; Bresnan 1982), and more specifically a *lexeme-based* approach in the sense of (Matthews 1974). As defined in (Fradin 2003) and references therein, the lexeme is considered to be an abstract entity, defined by its morphophonological properties, its meaning and its morphosyntactic category.<sup>1</sup> Concrete *forms* are built by inflectional morphology. We in particular follow the *Word and Paradigm* approach, e.g. Matthews, Aronoff and Stump (Matthews 1972; Zwicky 1985; Anderson 1992; Aronoff 1994; Stump 2001) and adopt a representation of forms based on the notions of *stems* and *exponents* as used by (Robins 1959) and (Matthews 1974): stems are what remains of a form once all exponents have been removed from it. In our approach, (concrete) forms are built from (abstract) lexemes through (*form*) *realisation rules* which are applied to them. Such rules are defined in inferential-realisation models such as *Paradigm Function Morphology (PFM)* (Stump 2001) or *Network Morphology* (Corbett/Fraser 1993).

The aim of this paper is to provide a means for assessing the notion of morphological canonicity through original measures developed within our new morphological framework  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ . In particular, we introduce original measures for non-canonical phenomena such as *heteroclisis*, *deponency*, *defectiveness* and *overabundance*.

### 2 CANONICAL INFLECTION

The concept of *canonical typology* can be traced back to Corbett as in (Corbett 2003). It represents an attempt to better understand what exactly differs from a hypothetical ideal *canonical* stage in the different occurrences of non-canonical phenomena. Note that in this approach, *canonical inflection* must not be mistaken for *prototypical inflection*. Canonical inflection is rare. It corresponds to an ideal state, seldom, if ever, met, but that constitutes a purely theoretical space from which deviant phenomena can be formally distinguished (Corbett 2007).

---

\* *Author's address:* Laboratoire de Linguistique Formelle, Université Paris Diderot, 175 rue du Chevaleret, 75013 Paris, France. Email: geraldine.walther@linguist.jussieu.fr

<sup>1</sup> In particular, we also consider morphosyntactic information such as *argument structure* to be specified within the lexicon.

Canonical inflection is supposed to represent complete regularity, as well as an ideal correspondence between form and function such that different forms can most efficiently be distinguished from each other. In particular, it is a notion that affects both the relation between the cells of a given lexeme's paradigm and the corresponding cells belonging to two different lexemes' paradigms. Canonical inflection is thus defined through the comparison of both the cells of one given lexeme and the lexemes themselves.

Each cell in a given paradigm canonically shares the lexeme's stem but varies in terms of exponence depending on the morphosyntactic feature structure<sup>2</sup> expressed by the given form. On the other hand, across lexemes, the stems will canonically vary from one paradigm to the other while the exponents will remain the same according to their expressed morphosyntactic feature structures. These relations are illustrated through Tables 1 and 2.

In this work, we consider an inflectional paradigm canonical if it satisfies the criteria given in Table 3 (Corbett 2007). To these criteria we add the ones in Table 4 that further define canonical paradigm shape.<sup>3</sup> As stated for example in (Corbett 2007), deviation from these criteria leads to non-canonical paradigmatic properties.

A paradigm is considered canonical if it matches above mentioned criteria. The more it deviates from these criteria, the less canonical the paradigm. However, existing work on canonicity does not provide quantitative means to assess the degree of canonicity of a lexeme's paradigm. Such quantitative measures are the new feature we propose in section 4.

We present measures for four types of non-canonical inflection phenomena, namely *deponency*, *heteroclis*, *defectiveness* and *overabundance*. These measures are computed within the inferential realisational model for inflectional morphology *PARSLI*.

In the following section, we first provide a short description of the major features of this model. A complete formal description can be found in Appendix B. Section 4 then goes on with presenting the measures of canonicity developed within *PARSLI* that allow for quantitatively assessing the canonicity of a given paradigm.

FEATURES	LEXEME 1		LEXEME 2		
1 <sup>st</sup> p. Sing	stem1	-ma	stem2	-ma	<div style="display: flex; align-items: center; gap: 10px;"> <div style="width: 15px; height: 15px; background-color: #cccccc; border: 1px solid #000;"></div> same           <div style="width: 15px; height: 15px; background-color: #e0e0e0; border: 1px solid #000;"></div> different         </div>
2 <sup>nd</sup> p. Sing	stem1	-sa	stem2	-sa	
3 <sup>rd</sup> p. Sing	stem1	-ta	stem2	-ta	
1 <sup>st</sup> p. Pl	stem1	-mo	stem2	-mo	
2 <sup>nd</sup> p. Pl	stem1	-so	stem2	-so	
3 <sup>rd</sup> p. Pl	stem1	-to	stem2	-to	

Table 1: Comparison over the cells of a given lexeme.

<sup>2</sup> Or what (Stump 2001) would refer to as morphosyntactic *property sets*.

<sup>3</sup> However, among the additional criteria, criterion 1 derives directly from criterion 2 in (Corbett 2007) and criterion 3 can be seen as derived from criterion 3 in (Corbett 2007).

FEATURES	LEXEME 1		LEXEME 2		
1 <sup>st</sup> p. Sing	stem1	<i>-ma</i>	stem2	<i>-ma</i>	<div style="display: flex; align-items: center; gap: 10px;"> <div style="width: 15px; height: 15px; background-color: #cccccc; border: 1px solid #000;"></div> same           <div style="width: 15px; height: 15px; background-color: #e0e0e0; border: 1px solid #000;"></div> different         </div>
2 <sup>nd</sup> p. Sing	stem1	<i>-sa</i>	stem2	<i>-sa</i>	
3 <sup>rd</sup> p. Sing	stem1	<i>-ia</i>	stem2	<i>-ia</i>	
1 <sup>st</sup> p. Pl	stem1	<i>-mo</i>	stem2	<i>-mo</i>	
2 <sup>nd</sup> p. Pl	stem1	<i>-so</i>	stem2	<i>-so</i>	
3 <sup>rd</sup> p. Pl	stem1	<i>-to</i>	stem2	<i>-to</i>	

Table 2: Comparison across lexemes.

	COMPARISON ACROSS CELLS OF A LEXEME	COMPARISON ACROSS LEXEMES
1 COMPOSITION/STRUCTURE	<i>same</i>	<i>same</i>
2 LEXICAL MATERIAL ( $\approx$ shape of stem)	<i>same</i>	<i>different</i>
3 INFLECTIONAL MATERIAL ( $\approx$ shape of inflection)	<i>different</i>	<i>same</i>
4 OUTCOME ( $\approx$ shape of inflected word)	<i>different</i>	<i>different</i>

Table 3: Criteria for Canonical Inflection according to (Corbett 2007).

	CANONICAL INFLECTION
1 STEMS AND FEATURES	<p><i>There is no “mismatch between form and function” (Baerman 2007).</i></p> <p><i>Each lexeme has exactly one stem that combines with a series of exponents.</i></p>
2 COMPLETENESS	<p><i>There exists exactly one form corresponding to the expression of a specific morphosyntactic feature structure.</i></p>
3 INFLECTION CLASS	<p><i>All forms of a lexeme are built from one single inflection class.</i></p>

Table 4: Additional criteria for Canonical Inflection.

### 3 INTRODUCING PARSLI

The name PARSLI stands for “PARadigm Shape and Lexicon Interface”. PARSLI is a formal model designed for representing morphological information stored within the (morphological) lexicon on the one hand and (morphological) grammar on the other and giving a description of each lexeme of a given language with regard to its own paradigm structure. It is the paradigm structure that accounts for the various non-canonical inflectional phenomena mentioned above.

#### 3.1 Defining the relevant notions

In PARSLI a lexeme is considered from the point of view of its formal participation in the inflectional process. Thus, we do not consider any specific semantics or possible derivational properties. In other words, we are here interested in the behaviour of what Fradin and Kerleroux refer to as *inflectemes* (Fradin/Kerleroux 2003),

as opposed to lexemes, and for which a (very) simplified definition could be "a lexeme minus its semantic and argument-structural information."

At this stage of its development,  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  does not make any claims about how exactly the realisation of the forms should be modelled. It solely focuses on the distribution of the morphological information between the morphological lexicon and the morphological grammar. The realisation of the forms by the realisation rules contained within the morphological grammar can be represented by any suitable independent inferential-realisation formalism.

### 3.2 Describing the $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ model of inflectional morphology

$\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  represents an inflecteme  $\mathbb{J}^4$  through seven defining elements:

1. the set of morphosyntactic feature structures  $\mathbb{J}$  can express,
2. the lexeme's morphosyntactic category,
3. an *inflection pattern*,
4. a *stem pattern*,
5. a *transfer rule* for stem selection,
6. a *transfer rule* for form realisation,
7. a pattern representing the paradigm

$\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  relies on the concept of *inflection class*. Note that the definition of an inflection class in  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  is not the traditional one, that is a particular paradigm type. In  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ , an inflection class is defined as a function associating morphosyntactic feature structures with corresponding realisation rules, i.e., a way to apply specific exponents corresponding to a given morphosyntactic feature structure. Each inflection class is partitioned into one or more *inflection zones* which are the core of  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ 's representation of inflection. As shown below, it is the selection of these inflection zones that determines an inflecteme's paradigm shape.

Inflection classes are the default associations of inflection zones that allow for computing the default paradigm structures of a language.<sup>5</sup> Using inflection zones from different inflection classes results in heterocclisis, as shown in Section 4 below.

Similarly,  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  also uses the concept of *stem class*, i.e., a function associating morphosyntactic feature structures with corresponding stem formation rules. Each stem class is partitioned into *stem zones*.

As will be shown below, these elements allow for the realisation of one of a given lexeme's form corresponding to a given morphosyntactic feature structure. In order

---

<sup>4</sup> I.e. the morphological part of a lexical entry.

<sup>5</sup> Usually this corresponds to the most frequent combination for a language's lexical items.



to illustrate the different steps of form realisation, we here outline the derivation of forms for the Italian adjectival inflecteme *CARO dear*. However, the complete process will be clearest after reading the formal definitions and illustrations in Appendixes B and C.

	MASC	FEM
SG	<i>karo</i>	<i>kara</i>
PL	<i>kari</i>	<i>kare</i>

Table 5: The paradigm of *CARO dear* in Italian.

The Italian inflecteme *CARO* can express four distinct morphosyntactic feature structures:

- {GENDER *masc*, NUMBER *sg*}
- {GENDER *masc*, NUMBER *pl*}
- {GENDER *fem*, NUMBER *sg*}
- {GENDER *fem*, NUMBER *pl*}

A traditional representation of that paradigm would be as in Table 5.

### Stem formation

1. Each inflecteme being associated with a specific stem pattern, this stem pattern selects one particular stem zone corresponding to the morphosyntactic feature structure that is to be expressed. This stem zone is further used to obtain the stem formation rule.
  - *For CARO there is only one unique zone, associated with all four possible morphosyntactic feature structures.*
2. The (stem formation) transfer rule associated with the inflecteme computes the morphosyntactic feature structure that should be given as an input to the computed stem formation rule, given the morphosyntactic feature structure that is meant to be expressed for this given inflecteme;
  - *In the case of the inflecteme CARO, the stem formation transfer rule is the identity function, i.e., its output equals its input.*
3. The transformed feature structure is then associated with a specific stem formation rule through the inflection zone computed at step 1 above;
  - *This stem formation rule is the same for all morphosyntactic feature structures applicable to the inflecteme CARO. It will always compute the same stem regardless the morphosyntactic feature given as an input.*

4. The stem formation rule computes the correct stem form for the inflecteme (this rule may be expressed formally with any suitable realisation based formalism).

- *The inflecteme CARO has only one possible stem [kar].*

## Inflection

1. In parallel, the inflection pattern associated with the inflecteme selects a specific inflection zone for the form realisation corresponding to a specific morphosyntactic feature structure;

- *Given an input feature [GENDER fem, NUMBER sg] this zone will also be the inflection zone associated with the plural forms of the feminine. Whether or not this zone is the same for the other forms of the paradigm depends on the general structure of the language.<sup>6</sup>*

2. The (form realisation) transfer rule associated with the inflecteme computes the morphosyntactic feature structure that should be given as an input to the computed realisation rule, given the morphosyntactic feature structure that is meant to be expressed for this given inflecteme;

- *In the case of the inflecteme CARO, the form realisation transfer rule is the identity function, i.e., its output equals its input, just as for the stem formation transfer rule.*

3. The transformed feature structure is then associated with a specific realisation rule through the inflection zone computed at step 1 above;

- *In the case of the inflecteme CARO for the input feature structure [GENDER fem, NUMBER sg], this form realisation rule specifies the adding of the feminine singular exponent [a] to the stem.*

## Form generation

1. Finally, the realisation rule obtained in *Inflection 3* is applied to the stem computed in *Stem formation 4* and the transformed morphosyntactic feature structure obtained in *Inflection 2*. It computes the correct form for a given input feature structure of the inflecteme. This realisation rule may be expressed formally with any suitable realisation based formalism.

- *In the case of the inflecteme CARO for the input feature [GENDER fem, NUMBER sg], the realised form is thus [kara].*

Note that transfer rules most often default to the identity function. Whenever they differ from the identity function, they express “a mismatch between form and function“ as (Baerman 2007) puts it. They are used for modelling deponency.

---

<sup>6</sup> For a more detailed representation thereof, see the representation of heterocclisis in Section 4.

Inflection zones used by a given inflecteme  $\mathcal{J}$ , i.e., the set of zones associated with it by its inflection rule, are called its *inflection pattern*. They build the inflecteme's paradigm. The set of  $\mathcal{J}$ 's stem zones is called its *stem pattern*.

Inflection classes are defined as *the most natural combination of inflection zones*, i.e., those that are used together by a majority of inflectemes. They are default inflection patterns.

Sometimes a given morphosyntactic feature structure can be associated with more than one stem zone by a stem pattern and more than one inflection rule by the inflection pattern. However, in such a case, nothing enforces that each stem zone can be equally combined with each inflection zone. The situation is even worse when transfer rules differ from the identity function.

Therefore, we need a way to express the possible combinations of stem zones and inflection zones: the combinations are what we call *subpatterns*. These subpatterns are 4-tuples consisting of a stem zone, an inflection zone and two transfer rules. They express the possible combinations for a given inflecteme. A subpattern requires that the sets of morphosyntactic feature structures associated with the two zones have a non-empty intersection. The set of a given inflecteme's subpatterns is the inflecteme's *pattern*.

In the following section, we will show how the measures developed within  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  allow for measuring the canonicity of paradigms in terms of *y deponency*, *heteroclisys*, *defectiveness* and *overabundance*.<sup>7</sup>

## 4 EXPRESSING AND MEASURING NON-CANONICAL PARADIGM SHAPES WITH $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$

In this section we present non-canonical phenomena affecting paradigm structures and the associated measures.

### 4.1 Stem alternations, allomorphy and suppletion

Suppletion comes in two types: *stem suppletion* and *form suppletion* (Boyé 2006). Stem suppletion occurs whenever, inside a paradigm, the forms' exponents remain regular, but their stems vary. This is for example the case for the French verb ALLER *to go* which has four different stems, *all-*, *v-*, *i-* and *aill-*. Form suppletion corresponds to cases where a whole form is inserted in a paradigm cell that should canonically be filled by a certain stem and the exponent corresponding to this specific cell. Form suppletion is described in (Bonami/Boyé 2002) for the French verbe ÊTRE *to be* in the present indicative. For this verb, the 1<sup>st</sup> person plural form *sommes*, for example, is unique in not using the regular 1<sup>st</sup> person plural exponent *-ons* that canonically appears with corresponding forms of other verbs (see Table 6).

---

<sup>7</sup> For a list of the symbols used in the more formal definitions, please refer to Appendix A.

	SINGULAR	PLURAL
p1	<i>suis</i>	<i>sommes</i>
p2	<i>es</i>	<i>êtes</i>
p3	<i>est</i>	<i>sont</i>

Table 6: Form suppletion in the present indicative paradigm of French *être* 'to be'

#### 4.1.1 Formal definition of allomorphy

Let  $\mathcal{I} = (\mathcal{K}_{\mathcal{I}}, C_{\mathcal{I}}, s_{\mathcal{I}}, f_{\mathcal{I}}, \mathcal{F}_{\psi_{\mathcal{I}}}, \mathcal{F}_{\chi_{\mathcal{I}}}, P_{\mathcal{I}})$  be an inflecteme, its stem pattern  $s_{\mathcal{I}}$  associates (at least) one stem zone  $\zeta_{s_{\mathcal{I}}, \kappa}$  to a morphosyntactic feature structure  $\kappa \in \mathcal{K}_{\mathcal{I}}$ .<sup>8</sup>

A stem selection rule  $s$  allows for formally representing morphomic (in the sense of (Aronoff 1994)) structures in stem selection, such as can be observed for Latin verbs.

#### 4.1.2 Example: Latin verb stems

In Latin, the distribution of the three existing stems available for all Latin verbs is morphomic in the sense that all verbs use the same stem pattern. This stem pattern is partitioned into three stem zones. Tables 7 and 8 give a schematic representation of the three stem zones.

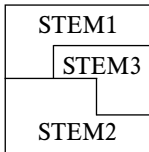


Table 7: Stem zones in the Latin active (sub-) paradigm

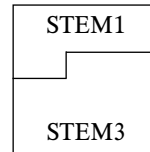


Table 8: Stem zones in the Latin passive (sub-) paradigm

STEM	ACT. SUBPARADIGM	PASS. SUBPARADIGM
STEM1	<i>imperf. finite</i>	<i>imperf. finite</i>
STEM2	<i>perf. finite</i>	
STEM3	<i>active future part.</i>	<i>passive past part.</i> <i>perf. finite (periphr.)</i>

Table 9: Morphomic combinations between morphosyntactic features and Latin verb stems

<sup>8</sup> Usually  $\zeta_{s_{\mathcal{I}}, \kappa}$  associates all compatible morphosyntactic feature structures with one unique stem formation rule.

## 4.2 Deponency

Some Croatian nouns use singular forms to express plural, as shown in the data presented in (Baerman 2006). This mismatch between form and function is what, following Baerman (Baerman 2007), we name *deponency*.

As shown in (Baerman 2006), Croatian nouns are inflected according to a number of different declension classes. Some classes that are relevant for our discussion are shown in Table 10. The data shows that the nouns *dete* 'child' and *tele* 'calf' inflect in the plural according to the singular pattern of respectively the A-STEM and I-STEM inflection classes. Using singular inflection to express the plural results in this mismatch between form and function.<sup>9</sup>

	(FEMININE) A-STEM		(FEMININE) I-STEM	
	<i>žena</i> 'woman'		<i>stvar</i> 'thing'	
	SINGULAR	PLURAL	SINGULAR	PLURAL
NOM	<i>žen-a</i>	<i>žen-e</i>	<i>stvar</i>	<i>stvar-i</i>
ACC	<i>žen-u</i>	<i>žen-e</i>	<i>stvar</i>	<i>stvar-i</i>
GEN	<i>žen-e</i>	<i>žen-a</i>	<i>stvar-i</i>	<i>stvar-i</i>
DAT	<i>žen-i</i>	<i>žen-ama</i>	<i>stvar-i</i>	<i>stvar-ima</i>
INS	<i>žen-om</i>	<i>žen-ama</i>	<i>stvar-i</i>	<i>stvar-im</i>

Table 10: Croatian noun inflection

	(FEMININE) A-STEM		(FEMININE) I-STEM	
	<i>žena</i> 'woman'		<i>stvar</i> 'thing'	
	SINGULAR	PLURAL	SINGULAR	PLURAL
NOM	<i>dete</i>	<i>deca</i>	<i>tele</i>	<i>telad</i>
ACC	<i>dete</i>	<i>decu</i>	<i>tele</i>	<i>telad</i>
GEN	<i>deteta</i>	<i>dece</i>	<i>teleta</i>	<i>telad</i>
DAT	<i>detetu</i>	<i>deci</i>	<i>teletu</i>	<i>teladi (ma)</i>
INS	<i>detetom</i>	<i>decom</i>	<i>teletom</i>	<i>teladi (ma)</i>

Table 11: Croatian deponent noun inflection

### 4.2.1 Formal definition of deponency

Let  $\mathcal{J}=(\mathcal{K}_{\mathcal{J}}, C_{\mathcal{J}}, s_{\mathcal{J}}, f_{\mathcal{J}}, \mathcal{F}_{\psi_{\mathcal{J}}}, \mathcal{F}_{\chi_{\mathcal{J}}}, P_{\mathcal{J}})$  be an inflecteme.

The “mismatch between form and function” stated by (Baerman, 2007) to be the definition of deponency occurs whenever the morphosyntactic features expressed by a given inflecteme's form  $f$  do not match the morphosyntactic features  $\varkappa$  usually expressed by the realisation rule  $\varphi$  used to build that form of this given inflecteme.

<sup>9</sup> For more data on deponency, the reader may refer to the large database put together by the Surrey Morphology Group: <http://www.smg.surrey.ac.uk/deponency>.

As mentioned above, within  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  this means that the transfer rule  $\mathcal{T}_{\mathcal{X}_{\mathcal{J}}}$  differs from the identity function, or, in other words, the morphosyntactic feature structure  $\mathcal{T}_{\mathcal{X}_{\mathcal{J}}}(\varkappa)$  expressed by  $f_{\mathcal{J}}$  differs from the morphosyntactic feature structure  $\varkappa$  that has been associated by the appropriate inflection zone  $\xi$  through  $f_{\mathcal{J}}$ .

More precisely, a given inflecteme  $\mathcal{J}$  is said to be *deponent* iff there exists at least one form  $f$  in its paradigm  $\mathfrak{P}$  built in way such that

$$\varkappa \neq \mathcal{T}_{\mathcal{X}_{\mathcal{J}}}(\varkappa).$$

Let  $\mathcal{P}_{\mathcal{X}_{\mathcal{J}}} = \{\mathcal{K}_{\mathcal{J},1}, \dots, \mathcal{K}_{\mathcal{J},n}\}$  be the smallest partition of  $\mathcal{X}_{\mathcal{J}}$  such that  $f_{\mathcal{J}}$  associates each  $\varkappa \in \mathcal{K}_{\mathcal{J},i}$  with the same  $\mathcal{K}_{\mathcal{J},i}$ .<sup>10</sup>

An inflecteme is considered to be *semi-deponent* iff for at least one element of  $\mathcal{K}_{\mathcal{J},i} \subset \mathcal{P}_{\mathcal{X}_{\mathcal{J}}}$  but not all, the restriction  $\mathcal{T}_{\mathcal{X}_{\mathcal{J}}|\mathcal{K}_{\mathcal{J},i}}$  of  $\mathcal{T}_{\mathcal{X}_{\mathcal{J}}}$  to  $\mathcal{K}_{\mathcal{J},i}$  is the identity function.

#### 4.2.2 Deponency Index

For the non-canonical phenomenon of deponency we can thus compute a measure of canonicity. We call this measure the *deponency index*. The deponency index  $\mathcal{D}_{\mathcal{J}}$  of an inflecteme  $\mathcal{J}$  is defined as the number of elements of the form  $\mathcal{K}_{\mathcal{J},i}$  in  $\mathcal{P}_{\mathcal{X}_{\mathcal{J}}}$  such that  $\mathcal{T}_{\mathcal{X}_{\mathcal{J}}|\mathcal{K}_{\mathcal{J},i}} \neq \text{id}$ :

$$\mathcal{D}_{\mathcal{J}} = |\{\mathcal{K}_{\mathcal{J},i} \in \mathcal{P}_{\mathcal{X}_{\mathcal{J}}} | \mathcal{T}_{\mathcal{X}_{\mathcal{J}}|\mathcal{K}_{\mathcal{J},i}} \neq \text{id}\}|$$

Hence, an inflecteme is deponent iff  $\mathcal{D}_{\mathcal{J}} > 0$ .

An inflecteme is semi-deponent iff  $|\mathcal{P}_{\mathcal{X}_{\mathcal{J}}}| > \mathcal{D}_{\mathcal{J}} > 0$ .

Conversely, a non-deponent inflecteme verifies  $\mathcal{D}_{\mathcal{J}} = 0$ .

#### 4.2.3 Example: Croatian nouns

The Croatian data presented above can be modelled within  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  with a transfer rule. Hence, an inflecteme building its paradigm as described above entails a transfer rule  $\mathcal{T}_{\mathcal{X}_{\mathcal{J}}}$  within its definition for which

$$\mathcal{T}_{\mathcal{X}_{\mathcal{J}}}(\{\text{NUMBER plural}\}) = \{\text{NUMBER singular}\}.$$

Hence it is a semi-deponent noun: the transfer rule differs from the identity function for the morphosyntactic feature structures containing the attribute-value pair  $\{\text{NUMBER plural}\}$ . For those containing  $\{\text{NUMBER singular}\}$ ,  $\mathcal{T}_{\mathcal{X}_{\mathcal{J}}} = \text{id}$ .

In other words, for both these two nouns, the smallest partition of  $\mathcal{X}_{\mathcal{J}}$  such that  $f_{\mathcal{J}}$  associates each  $\varkappa \in \mathcal{K}_{\mathcal{J},i}$  with the same  $\mathcal{K}_{\mathcal{J},i}$  consists of two subsets of  $\mathcal{X}_{\mathcal{J}}$ , one for singular feature structures (for which no deponency occurs) and one for plural ones.

---

<sup>10</sup> In the case of overabundant inflectemes, this does not concern a unique zone but the same set of zones.

Thus, the deponency index for these nouns is equal to 1. Since the total number of elements of that partition is 2, these lexemes are semi-deponents.

MASCULINE ANIMATE <i>CHLAP</i> 'boy'		MASCULINE INANIMATE <i>DUB</i> 'oak'		MASCULINE HETEROCLITE <i>OROL</i> 'eagle'	
SINGULAR	PLURAL	SINGULAR	PLURAL	SINGULAR	PLURAL
NOM <i>chlap</i>	<i>chlap-i</i>	<i>dub</i>	<i>dub-y</i>	<i>orol</i>	<i>orl-y</i>
GEN <i>chlap-a</i>	<i>chlap-ov</i>	<i>dub-a</i>	<i>dub-ov</i>	<i>orl-a</i>	<i>orl-ov</i>
DAT <i>chlap-ovi</i>	<i>chlap-om</i>	<i>dub-u</i>	<i>dub-om</i>	<i>orl-ovi</i>	<i>orl-om</i>
ACC <i>chlap-a</i>	<i>chlap-ov</i>	<i>dub</i>	<i>dub-y</i>	<i>orl-a</i>	<i>orl-y</i>
LOC <i>chlap-ovi</i>	<i>chlap-och</i>	<i>dub-e</i>	<i>dub-och</i>	<i>orl-ovi</i>	<i>orl-och</i>
INS <i>chlap-om</i>	<i>chlap-mi</i>	<i>dub-om</i>	<i>dub-mi</i>	<i>orl-om</i>	<i>orl-ami</i>

Table 12: Heterocclisis in Slovak masculine animal names inflection

### 4.3 Heterocclisis

Heterocclisis refers to the phenomenon where a lexeme's paradigm is built out of (at least) two, otherwise separate, inflection classes.

Examples of heterocclisis are (some) Slovak animal nouns. Indeed, in Slovak, most masculine animal nouns are inflected as masculine animate nouns in the singular, whereas they may (and for some lexemes, must) inflect as masculine inanimate nouns in the plural (except in specific cases, such as personification, which triggers the animate inflection even for plural forms) (Zauner 1973). Compare for example the inflection of *CHLAP* 'boy', *DUB* 'oak' and *OROL* 'eagle' in Table 12.<sup>11</sup>

#### 4.3.1 Definition

If all inflection zones associated with a given inflecteme  $\mathcal{J}$  belong to the same inflection class  $F$ , the inflecteme is canonical in the dimension of inflection class constitution. Conversely, if its inflection zones belong to at least two distinct inflection classes, the inflecteme is said to be *heterocclite*.

#### 4.3.2 Heterocclity Index

We define an inflecteme's *heterocclity index*  $\mathcal{H}_{\mathcal{J}}$  as the number of zones used to build an inflecteme's paradigm that are partitions of distinct inflection classes. In other words, this represents the number of inflection classes involved in the building of that inflecteme's paradigm.

More precisely,  $\mathcal{H}_{\mathcal{J}}$  is defined in the following way:

$$\mathcal{H}_{\mathcal{J}} = |\{\tilde{\xi} \mid \xi \in X_{\mathcal{J}}\}| - 1$$

<sup>11</sup> Both *CHLAP* and *DUB* have a regular inflection: *CHLAP* belongs to the standard inflection class for masculine animate stems ending with a consonant, whereas *DUB* belongs to the standard inflection class for masculine inanimate stems ending with what is called a hard or neutral consonant in the Slavic linguistic tradition.

where  $\tilde{\xi}$  stands for the inflection class to whose partition  $\xi$  belongs, and  $X_{\mathcal{J}}$  for the set of zones associated with  $\mathcal{J}$  through its inflection pattern.

Thus,  $\mathcal{J}$  is heteroclite iff  $\mathcal{H}_{\mathcal{J}} > 0$ .

### 4.3.3 Example 1: Slovak animal nouns

For the Slovak animal nouns described in Table 12, the inflection zone used for building the singular forms of the noun *OROL* 'eagle' is an element of the partition of the inflection class associated with animate nouns such as *CHLAP* 'boy', while the inflection zone used for the plural forms of such animal nouns belongs to the partition of the inflection class associated with inanimates like *DUB* 'oak'.

The heteroclicity index of such nouns is

$$\mathcal{H}_{OROL} = |\{\tilde{\xi} | \xi \in X_{OROL}\}| - 1 = |\tilde{\xi}_{F_{animate}, pl}| + |\tilde{\xi}_{F_{animate}, sg}| - 1 = 1 + 1 - 1 = 1$$

### 4.3.4 Example 2: Croatian nouns

Similarly, the Croatian nouns from Table 11 show inflection patterns producing the inflection zones listed in Table 13. These tables show that the corresponding Croatian nouns are not only deponent, but also heteroclite.

Thus, several-canonical phenomena may sometimes occur simultaneously in non-canonical paradigms.

INFLECTION CLASS	A: NEUTER STEM IN -ET	B: (FEMININE) STEM IN -A	C: (FEMININE) STEM IN -I
<i>DETE</i> 'child'	SG: $\xi_{A,sg}$	PL: $\xi_{B,sg}$	
<i>TELE</i> 'veal'	SG: $\xi_{A,sg}$		PL: $\xi_{C,sg}$

Table 13: Nominal inflection of Croatian heteroclite nouns

## 4.4 DEFECTIVENESS

Defectiveness (Baerman et al. 2010) refers to lexemes which display empty (missing) cells in their paradigm. Sometimes languages contain lexemes for which expected forms are simply unexisting; native speakers would always try avoiding having to build the corresponding forms. This is for example what we can observe with some French verbs such as *PAÎTRE* *to graze* for which there are no past tense forms available apart from the imperfect. Another example are the *pluralia tantum* described below.

### 4.4.1 Formal definition

A paradigm is considered *defective* iff there is at least one morphosyntactic feature structure  $\varkappa$  belonging to the set  $\mathcal{K}_{C_{\mathcal{J}}}$  of the morphosyntactic feature structures of the category of an inflecteme  $\mathcal{J}$  which  $f_{\mathcal{J}}$  does not associate with any inflectional zone  $\xi$ .

One can also say that an inflecteme  $\mathcal{J}$  is defective iff the set  $\mathcal{K}_{\mathcal{J}}$  of its morphosyntactic feature structures does not cover the set  $\mathcal{K}_{C_{\mathcal{J}}}$  of the morphosyntactic feature structures of its category.

$$\mathcal{K}_{\mathcal{J}} \subsetneq \mathcal{K}_{C_{\mathcal{J}}}$$



**Example: Pluralia Tantum** Another example are the nouns called *pluralia tantum* which only exist in the plural, cf. English TROUSERS, French VIVRES *food supplies* or Slovak VIANOCE *Christmas*.

Let us take the example of the French *pluralium tantum*  $\mathfrak{J}$  VIVRES food supplies. If we only consider the number features, we get the following defined morphosyntactic feature structures:

$$\mathcal{H}_{\mathfrak{J}} = [\text{NUMBER } plural]$$

while

$$\mathcal{H}_{C_{\mathfrak{J}}} = \mathcal{H}_{\text{nom}} = [\text{NUMBER } singular, \text{NUMBER } plural]$$

and

$$\mathcal{H}_{\mathfrak{J}} \subsetneq \mathcal{H}_{C_{\mathfrak{J}}}$$

#### 4.5 Overabundance

The obvious counterpart to defectiveness is the concept of *overabundance*. Overabundance occurs when cells of a paradigm contain more than one form. The notion has been introduced by Thornton and is discussed in (Thornton 2010) for Italian. Canonical overabundance characterises the case where *cell mates* of one given cell compete, without any morphological feature<sup>12</sup> permitting to choose one over the other. Table 14 shows examples thereof for Italian verbs.

	CELL-MATE 1	CELL-MATE 2
'languish' 3PL.PRS.SUBJ	<i>languano</i>	<i>languiscano</i>
'possess' 3PL.PRS.SUBJ	<i>possiedano</i>	<i>posseggano</i>
'possess' 3SG.PRS.SUBJ	<i>possieda</i>	<i>possegga</i>
'possess' 1SG.PRS.SUBJ	<i>possiedo</i>	<i>posseggo</i>

Table 14: Overabundance in Italian (Thornton, 2010)

In French, an example is given by the verb *ASSEOIR* 'to sit' that has two different forms in most cells as shown in Table 15.<sup>13</sup> All French verbs in *-ayer* also exhibit systematic overabundance (see Table 16). Indeed, for some cells, these verbs may use two competing stems (in *-ay-* and in *-ai-*) and therefore have two different inflected forms, which are morphologically equivalent (although semantic, pragmatic, sociolinguistic and other constraints may interfere).

<sup>12</sup> Or any other type of feature.

<sup>13</sup> See for example (Bonami/Boyé 2010) for a longer discussion thereof.

	IND.PRES	SINGULAR	PLURAL
p1		<i>assois</i>	<i>asseyons</i>
		<i>assieds</i>	<i>asseyons</i>
p2		<i>assois</i>	<i>asseyez</i>
		<i>assieds</i>	<i>asseyez</i>
p3		<i>assoit</i>	<i>assoient</i>
		<i>assied</i>	<i>asseyent</i>

Table 15: Overabundance in French *asseoir* 'to sit'

	IND.PRES	SINGULAR	PLURAL
p1		<i>balaye</i>	<i>balayons</i>
		<i>balaie</i>	<i>balayons</i>
p2		<i>balayes</i>	<i>balayez</i>
		<i>balaies</i>	<i>balayez</i>
p3		<i>balaye</i>	<i>balayent</i>
		<i>balaie</i>	<i>balaient</i>

Table 16: Overabundance in French *balayer* 'to sweep'

#### 4.5.1 Formal definition

A paradigm is considered *overabundant* iff there is at least one morphosyntactic feature structure  $\varkappa$  belonging to the set  $\mathcal{K}_{C_{\mathcal{J}}}$  of the morphosyntactic feature structures of the category of an inflecteme  $\mathcal{J}$  which  $f_{\mathcal{J}}$  associates with more than one inflection zone. In that case,  $f_{\mathcal{J}}$  is a generic binary relation and not a function.

$$f_{\mathcal{J}}(\varkappa) = S, \text{ where } |S| > 1.$$

**Example: Italian overabundant verbs** Table 14 shows examples of overabundant Italian verbs.<sup>14</sup>

In this case, the inflecteme LANGUIRE has a inflection pattern  $f_{\text{LANGUIRE}}$  which associates the morphosyntactic feature structure  $\varkappa = \{3\text{PL.PRS.SUBJ}\}$  with two inflection zones, each producing a different realisation rule  $\varphi_1$  and  $\varphi_2$ . These two rules thus give rise to two distinct forms within the paradigm  $\mathfrak{B}_{\text{LANGUIRE}}$  expressing  $\varkappa$ :  $f_1 = \text{languano}$  and  $f_2 = \text{languiscano}$ .

#### 4.6 Canonical Inflection

From the definitions of non-canonical phenomena above, we can deduce the following definition of *Canonical Inflection*.

Canonical inflection corresponds to the case where the inflection pattern  $f_{\mathcal{J}}$  of an inflecteme  $\mathcal{J}$  associates the morphosyntactic feature structures belonging to the set of morphosyntactic feature structures  $\mathcal{K}_{\mathcal{J}}$  for which  $\mathcal{J}$  is defined with inflection zones that constitute the complete set of elements contained within the partition of one unique inflection class  $F$ .

In particular, this entails that for all morphosyntactic feature structures  $\varkappa$ , the inflection pattern  $f_{\mathcal{J}}$  associates  $\varkappa$  with one unique element of the partition of  $F$ .

Moreover, the stem pattern associates every  $\varkappa \in \mathcal{K}_{\mathcal{J}}$  with a stem zone  $\zeta$  belonging to a stem class  $\Gamma$  containing only this one stem zone  $\zeta$  and that produces a unique stem formation rule  $\sigma$ , whatever  $\varkappa$ . In other words,  $\mathcal{J}$  has a unique stem.

<sup>14</sup> The data is borrowed from (Thornton 2010).

Finally, the transfer rule  $\mathcal{T}_{\chi_{\mathcal{J}}}$  is the identity function and the set of morphosyntactic feature structures  $\mathcal{X}_{\mathcal{J}}$  defined for  $\mathcal{J}$  equals the set of morphosyntactic feature structures  $\mathcal{X}_{C_{\mathcal{J}}}$  defined for  $\mathcal{J}$ 's morphosyntactic category  $C_{\mathcal{J}} \in \mathcal{C}$ .

The same holds for the transfer rule  $\mathcal{T}_{\psi_{\mathcal{J}}}$ .

## 5 CONCLUSION

We have presented  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}\mathbb{I}$ , a formal model of inflectional morphology.  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}\mathbb{I}$  being completely formalised, it can be implemented. Such an implementation would allow for the comparison of complete morphological descriptions with regard to their complexity. Indeed, previous experiments on complexity evaluation with  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}\mathbb{I}$  and its implementation within the Alexina lexical framework (Sagot 2010) have already been conducted (Sagot/Walther 2011). The usefulness of  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}\mathbb{I}$  to build morphological descriptions with reduced descriptive complexity has also been shown in (Walther/Sagot 2011).

But most importantly, in the domain of *Canonical Typology*,  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}\mathbb{I}$  contains original measures that allow for quantitatively assessing the canonicity of paradigms in the sense of the qualitative characterisation proposed by the approaches developed within *Canonical Typology* (Corbett 2003).

## A SUMMARY OF THE NOTATIONS IN $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}\mathbb{I}$

We use the following notations in the formal definitions:

- A morphosyntactic feature structure will be noted  $\alpha$ ,
- an inflection rule  $f$ ,
- an inflection class  $F$ ,
- an inflection zone  $\xi$ ,
- a stem selection rule  $s$ ,
- a stem class  $\Gamma$ ,
- a stem zone  $\zeta$ ,
- a transfer rule  $\mathcal{T}$ 
  - a stem transfer rule  $\mathcal{T}_{\psi_{\mathcal{J}}}$ ,
  - an inflection transfer rule  $\mathcal{T}_{\chi_{\mathcal{J}}}$ ,
- and a pattern  $P$ .

## B A FORMAL DEFINITIONS OF THE $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}\mathbb{I}$ MODEL

The next two appendixes provide the actual formalisation underlying the  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}\mathbb{I}$  model and a formalised representation of paradigm building within the model.

### B.1 Phonological material

An elementary sequence of phonological material  $e$  is a segmental or suprasegmental combination of sounds. The set of all elementary sequences of phonological material is noted  $E$ .

## B.2 Morphosyntactic features

### B.2.1 Features and feature structures

In this document, we define a *morphosyntactic feature structure*  $\varkappa$  as a set of attribute-value pairs. PARSLI makes no strong assumptions about how the feature structures are organised with regard to one another.

### B.2.2 Morphosyntactic categories

The set of all feature structures used in a given complete morphological description of an inflecteme is noted  $\mathcal{H}$ .

An *inflecteme*  $\mathcal{I}$  will be assigned a category depending on the feature structure it has information about: an inflecteme  $\mathcal{I}$  from a *category*  $\mathcal{C}$  will cover a subset  $\mathcal{H}_{\mathcal{I}}$  of the morphosyntactic feature structure set  $\mathcal{H}_{\mathcal{C}} \subset \mathcal{H}$  specific to that category.

The set of all categories is noted  $\mathcal{C}$ .

## B.3 Stems

### B.3.1 Definition

A *stem*  $r$  is an elementary sequence of phonological material. The set of all stems is noted  $R$ .

$$r \in R \subset E$$

### B.3.2 Stem formation rule

Stem formation is expressed through stem formation rules. A stem formation rule  $\sigma$  is a function from  $\mathcal{H}$  to  $E$  which takes a specific morphosyntactic feature structure  $\varkappa \in \mathcal{H}$  as an input so as to produce a phonological material  $e'$  expressing that feature.

$$\begin{aligned} \sigma : \mathcal{H} &\rightarrow E \\ \sigma(\varkappa) &= e' \end{aligned}$$

The set of all stem formation rules is noted  $\Sigma$ .

### B.3.3 Stem class

A stem class  $\Gamma$  is a function from  $\mathcal{H}_{\Gamma} \subset \mathcal{H}$  to  $\Sigma$ .

### B.3.4 Stem zones

Let  $\Gamma$  be a stem class defined over a set  $\mathcal{H}_{\Gamma}$  of morphosyntactic feature structures. For each  $\Gamma$  a unique partition of  $\mathcal{H}_{\Gamma}$  is defined, whose members are noted  $\mathcal{H}_{\Gamma,k}$ , such that:<sup>15</sup>

$$\mathcal{H}_{\Gamma} = \bigsqcup_k \mathcal{H}_{\Gamma,k}$$

---

<sup>15</sup> " $\bigsqcup$ " denotes the union of disjoint sets.

A stem zone  $\zeta$  for  $\Gamma$  is then defined as a pair

$$\zeta = (\mathcal{K}_{\Gamma,k}, \Gamma)$$

where  $\mathcal{K}_{\Gamma,k}$  is one element of the partition.

Let  $\zeta = (\mathcal{K}_{\Gamma,k}, \Gamma)$  be a zone for  $\Gamma$ . We define the operators  $\sim$  and  $\hat{\phantom{x}}$  as follows:  $\tilde{\zeta}$  is the second element of  $\zeta$ , i.e., its stem class  $\Gamma$ , and  $\hat{\zeta}$  is the first element of  $\zeta$ , i.e., the corresponding element of the partition of  $\mathcal{K}_\Gamma$ .

The set of zones for a stem class  $\Gamma$  is noted  $Z(\Gamma)$ . The set of all stem zones for all stem classes is noted  $\mathcal{Z}$ .

$$\mathcal{Z} = \bigcup_{\Gamma} Z(\Gamma)$$

### B.3.5 Stem pattern

A *stem pattern* is a binary relation  $s$  associating an element from a given  $\mathcal{K}_s \subset \mathcal{K}$  with one or more stem zones. A given morphosyntactic feature structure  $\varkappa \in \mathcal{K}_f$  will be associated through  $s$  with stem zones of the form  $\zeta = (\mathcal{K}_{\Gamma,k}, \Gamma)$ . From there we can retrieve the stem formation rule  $\sigma \in \Sigma$  corresponding to a given  $\varkappa \in \mathcal{K}_s$ :

$$\text{If } (\varkappa, \zeta) \in s \wedge \zeta = (\mathcal{K}_{\Gamma,k}, \Gamma),$$

then, provided we are given a certain  $\varkappa' \in \mathcal{K}_{\Gamma,k}$  (be it equal to  $\varkappa$  or not), one of the corresponding stem formation rule  $\sigma$  verifies

$$\sigma = \tilde{\zeta}(\varkappa') = \Gamma(\varkappa')$$

## B.4 Inflection

### B.4.1 Realisation rule

Inflection is expressed through realisation rules. A realisation rule  $\varphi$  is a function from  $E \times \mathcal{K}$  to  $E$  which takes specific phonological material  $e^{16}$  as an input so as to produce a modified phonological material  $e'$  in order to express a specific morphosyntactic feature structure  $\varkappa \in \mathcal{K}$ .

$$\begin{aligned} \varphi : E \times \mathcal{K} &\rightarrow E \\ \varphi(e, \varkappa) &= e' \end{aligned}$$

The set of all realisation rules is noted  $\Phi$ .

### B.4.2 Inflection class

An inflection class  $F$  is a function from  $\mathcal{K}_F \subset \mathcal{K}$  to  $\Phi$ .

---

<sup>16</sup> Namely the stem produced by the corresponding stem formation rule.

### B.4.3 Inflection zones

Let  $F$  be an inflection class defined over a set  $\mathcal{K}_F$  of morphosyntactic feature structures.

For each  $F$  is defined a unique partition of  $\mathcal{K}_F$ , whose members are noted  $\mathcal{K}_{F,k}$ , such that:

$$\mathcal{K}_F = \bigsqcup_k \mathcal{K}_{F,k}$$

An inflection zone  $\xi$  for  $F$  is then defined as a pair

$$\xi = (\mathcal{K}_{F,k}, F)$$

where  $\mathcal{K}_{F,k}$  is one element of the partition.

Let  $\xi = (\mathcal{K}_{F,k}, F)$  be a zone for  $F$ . We define the operators  $\tilde{\phantom{x}}$  and  $\hat{\phantom{x}}$  as follows:  $\tilde{\xi}$  is the second element of  $\xi$ , i.e., its inflection class  $F$ , and  $\hat{\xi}$  is the first element of  $\xi$ , i.e., the corresponding element of the partition of  $\mathcal{K}_F$ .

The set of zones for an inflection class  $F$  is noted  $X(F)$ . The set of all inflection zones for all inflection classes is noted  $\mathcal{X}$ .

$$\mathcal{X} = \bigcup_F X(F)$$

### B.4.4 Inflection pattern

An *inflection pattern* is a binary relation  $f$  associating an element from a given  $\mathcal{K}_f \subset \mathcal{K}$  with one or more inflection zones. A given morphosyntactic feature structure  $\varkappa \in \mathcal{K}_f$  will be associated through  $f$  with inflection zones of the form  $\zeta = (\mathcal{K}_{F,k}, F)$ . From there we can retrieve the inflectional function  $\varphi \in \Phi$  corresponding to a given  $\varkappa \in \mathcal{K}_f$ :

$$\text{If } (\varkappa, \xi) \in f \wedge \xi = (\mathcal{K}_{F,k}, F)$$

then, provided we are given a certain  $\varkappa' \in \mathcal{K}_{F,k}$  (be it equal to  $\varkappa$  or not), one of the corresponding inflectional function  $\varphi$  verifies

$$\varphi = \tilde{\xi}(\varkappa') = F(\varkappa')$$

### B.5 Transfer rules

We define a transfer rule  $\mathcal{T}$  as a function from its domain  $\mathcal{K}_{\mathcal{T}} \in \mathcal{K}$  to  $\mathcal{K}$ .

Given an inflecteme  $\mathcal{I}$ , there are two types of transfer rules. One  $\mathcal{T}_{\psi_{\mathcal{I}}}$  for stem formation and one  $\mathcal{T}_{\chi_{\mathcal{I}}}$  for inflection.

## B.6 Pattern

### B.6.1 Subpattern

A subpattern is defined for a given inflecteme  $\mathfrak{I}$ . It is a 4-tuple consisting of a stem zone  $\zeta$ , an inflection zone  $\xi$  and two transfer rules,  $\mathcal{T}_{\psi_{\mathfrak{I}}}$  and  $\mathcal{T}_{\chi_{\mathfrak{I}}}$ . To be valid, a subpattern requires that the set of morphosyntactic feature structures  $\hat{\zeta} \in \mathcal{K}$  and  $\hat{\xi} \in \mathcal{K}$  associated respectively with  $\zeta$  and  $\xi$  have a non-empty intersection.

$$\hat{\zeta} \cap \hat{\xi} \neq \emptyset$$

### B.6.2 Pattern

A pattern  $P$  is the set of all valid subpatterns defined for a given inflecteme  $\mathfrak{I}$ .

## B.7 Inflectemes

### B.7.1 Definition

#### Formal definition of an inflecteme:

An inflecteme  $\mathfrak{I}$  is a 7-tuple  $(\mathcal{K}_{\mathfrak{I}}, \mathcal{C}_{\mathfrak{I}}, s_{\mathfrak{I}}, f_{\mathfrak{I}}, \mathcal{T}_{\psi_{\mathfrak{I}}}, \mathcal{T}_{\chi_{\mathfrak{I}}}, P_{\mathfrak{I}})$ , where

- $\mathcal{K}_{\mathfrak{I}}$  is the set of morphosyntactic features  $\varkappa$  expressible by  $\mathfrak{I}$ ,
- $\mathcal{C}_{\mathfrak{I}}$  is  $\mathfrak{I}$  morphosyntactic category, and  $\mathcal{C}_{\mathfrak{I}} \in \mathcal{C}$ , where  $\mathcal{C}$  is the set of morphosyntactic categories that exist in a morphological description for a given language,
- $s_{\mathfrak{I}}$  is a *stem pattern*, a binary relation from  $\mathcal{K}_{\mathfrak{I}}$  to  $\mathcal{Z}_{S_{\mathfrak{I}}}$ , the set of *stem zones* compatible with  $\mathfrak{I}$ ;  $\mathcal{Z}_{S_{\mathfrak{I}}} \subset \mathcal{Z}$ , where  $\mathcal{Z}$  is the set of all stem zones in a morphological description of a given language,
- $f_{\mathfrak{I}}$  is a *inflection pattern*, binary relation from  $\mathcal{K}_{\mathfrak{I}}$  to  $\mathcal{X}_{f_{\mathfrak{I}}}$ , the set of *inflection zones* according to which a given inflecteme is inflected;  $\mathcal{X}_{f_{\mathfrak{I}}} \subset \mathcal{X}$ , where  $\mathcal{X}$  is the set of all inflection zones in a morphological description of a given language,
- $\mathcal{T}_{\psi_{\mathfrak{I}}}$  is a *transfer rule*, i.e., a function defined over at least all morphosyntactic feature structures  $\varkappa \in \mathcal{K}_{\mathfrak{I}}$ , such that  $\mathcal{T}_{\psi_{\mathfrak{I}}}(\varkappa)$  belongs to the set of morphosyntactic features realised through the stem zones defined for  $\mathcal{K}_{\mathfrak{I}}$ .
- $\mathcal{T}_{\chi_{\mathfrak{I}}}$  is a *transfer rule*, i.e., a function defined over at least all morphosyntactic feature structures  $\varkappa \in \mathcal{K}_{\mathfrak{I}}$ , such that  $\mathcal{T}_{\chi_{\mathfrak{I}}}(\varkappa)$  belongs to the set of morphosyntactic features realised through the inflection zones defined for  $\mathcal{K}_{\mathfrak{I}}$ ;
- $P_{\mathfrak{I}}$  is a pattern, i.e., a set of subpatterns defined as a 4-tuple of the form  $(\zeta_{\varkappa}, \xi_{\varkappa}, \mathcal{T}_{\psi_{\mathfrak{I}}}, \mathcal{T}_{\chi_{\mathfrak{I}}})$ , where  $\zeta_{\varkappa}$  is a stem zone associated with a given morphosyntactic feature structure  $\varkappa$  through  $s_{\mathfrak{I}}$  and  $\xi_{\varkappa}$  an inflection zone associated to  $\varkappa$  through  $f_{\mathfrak{I}}$ .

## B.8 Paradigms

Let  $(\mathcal{K}_J, C_J, s_J, f_J, \mathcal{T}_{\psi_J}, \mathcal{T}_{\chi_J}, P_J)$  be an inflecteme.

### B.8.1 Forms

A form  $f$  is a combination of elementary sequences of phonological material. It expresses a set of morphosyntactic features  $\varkappa$  for the inflecteme  $J = (\mathcal{K}_J, C_J, s_J, f_J, \mathcal{T}_{\psi_J}, \mathcal{T}_{\chi_J}, P_J)$  and is obtained from a stem  $r$  of  $J$  by the realisation rule  $\varphi$  corresponding to one of the appropriate inflection zones  $\xi$ , obtained through the inflection pattern  $f_J$ .  $\varphi$  is then equal to  $\tilde{\xi}$ .

$$f = \varphi(r_J, \varkappa')$$

where  $\varkappa'$  is the output of  $\mathcal{T}_{\chi_J}(\varkappa)$ .

From there, we can also express  $f$  in the following way:

$$f = \tilde{\xi}(r_J, \mathcal{T}_{\chi_J}(\varkappa)) = \tilde{\xi}(\tilde{\zeta}(\mathcal{T}_{\psi_J}(\varkappa), \mathcal{T}_{\chi_J}(\varkappa)))$$

### B.8.2 Definition

A *paradigm*  $\mathfrak{P}_J$  of a given inflecteme  $J$  is the set of all form-morphosyntactic feature structure pairs  $(f, \varkappa)$  such that  $\varkappa \in \mathcal{K}_J$  and

$$f = \tilde{\xi}(\tilde{\zeta}(\mathcal{T}_{\psi_J}(\varkappa), \mathcal{T}_{\chi_J}(\varkappa)))$$

### B.8.3 Formal definition of canonical inflection

From the definitions of non-canonical phenomena above, we can deduce the following definition of *Canonical Inflection*.

#### Definition of Canonical Inflection

$\exists F$ ; such that  $\forall \varkappa \in \mathcal{K}_J, \tilde{f}_J(\varkappa, F)$  which means that  $|\{\tilde{\xi} | \xi \in \chi_{f_J}\}| - 1 = 0$

$\exists \Gamma$ , such that  $\forall \varkappa \in \mathcal{K}_J, \tilde{s}_J(\varkappa, \Gamma)$ ; where  $\Gamma$  is a function independant from  $\varkappa$

and  $\mathcal{T}_{\chi_J} = \text{id}$

and  $\mathcal{T}_{\psi_J} = \text{id}$

and  $\mathcal{K}_J = \mathcal{K}_{C_J}$ .

## C BUILDING A PARADIGM WITH PARSL

In this section we give a short example of how PARSL can be used to model the building of a given inflecteme's paradigm. As an illustration we shall use the simple case of an Italian adjectival paradigm, the paradigm of the inflecteme *CARO dear*.



### C.1 Definition of the inflecteme CARO within the lexicon

The inflecteme CARO is defined within the lexicon as the 7-tuple

$$\text{CARO} = (\mathcal{H}_{\text{CARO}}, \mathcal{C}_{\text{CARO}}, \mathcal{S}_{\text{CARO}}, \mathcal{f}_{\text{CARO}}, \mathcal{T}\psi_{\text{CARO}}, \mathcal{T}\chi_{\text{CARO}}, P_{\text{CARO}}),$$

where

$$\mathcal{H}_{\text{CARO}} = ([\text{GENDER } \textit{masc}, \text{NUMBER } \textit{sg}], [\text{GENDER } \textit{masc}, \text{NUMBER } \textit{pl}], \\ [\text{GENDER } \textit{fem}, \text{NUMBER } \textit{sg}], [\text{GENDER } \textit{fem}, \text{NUMBER } \textit{pl}], )$$

and the inflecteme's morphosyntactic category is

$$\mathcal{C}_{\text{CARO}} = \text{adjective}$$

Let us note  $\zeta$  the unique stem zone used for the building of this Italian adjective form. The stem pattern  $\mathcal{S}_{\text{CARO}}$  of CARO associates each possible morphosyntactic feature structure defined for CARO with this unique stem zone  $\zeta$ .

$$\mathcal{S}_{\text{CARO}} = ([[\text{GENDER } \textit{masc}, \text{NUMBER } \textit{sg}], \zeta], [[\text{GENDER } \textit{masc}, \text{NUMBER } \textit{pl}], \zeta], \\ [[\text{GENDER } \textit{fem}, \text{NUMBER } \textit{sg}], \zeta], [[\text{GENDER } \textit{fem}, \text{NUMBER } \textit{pl}], \zeta])$$

Let us note  $\xi_{\textit{masc}}$  and  $\xi_{\textit{fem}}$  the inflection zones used for the building of Italian adjective forms. The inflection pattern  $\mathcal{f}_{\text{CARO}}$  of CARO associates any morphosyntactic feature structure defined for CARO with either of these two inflection zones, depending on the corresponding gender feature.<sup>17</sup>

$$\mathcal{f}_{\text{CARO}} = ([[\text{GENDER } \textit{masc}], \xi_{\textit{masc}}], [[\text{GENDER } \textit{fem}], \xi_{\textit{fem}}])$$

The inflecteme CARO does not display form-function mismatches. Its transfer rules hence equal the identity function.

$$\mathcal{T}\psi_{\text{CARO}} = \text{id} \\ \mathcal{T}\chi_{\text{CARO}} = \text{id}$$

Having computed all the necessary elements we can now express the inflecteme's pattern  $P_{\text{CARO}}$ .

$$P_{\text{CARO}} = ([(\zeta, \xi_{\textit{masc}}, \mathcal{T}\psi_{\text{CARO}}, \mathcal{T}\chi_{\text{CARO}}), (\zeta, \xi_{\textit{fem}}, \mathcal{T}\psi_{\text{CARO}}, \mathcal{T}\chi_{\text{CARO}})])$$

---

<sup>17</sup> Stating the existence of two inflection zones for Italian adjectives has been decided on the properties of some Italian nouns. It is however clear that this is a descriptorial choice made by the author and that other representations would be possible as well.

## C.2 Building the paradigm of the inflecteme CARO

The stem zone  $\zeta$  is the unique element of the default stem class for Italian adjectives. It associates each morphosyntactic feature structure within  $\mathcal{K}_{\text{CARO}}$  with a unique stem formation rule  $\sigma$ . Hence we can compute the stem formation rule for the inflecteme CARO.

$$\forall \mathcal{x} \in \mathcal{K}_{\text{CARO}}, \sigma(\mathcal{x}) = r_{\text{CARO}} \text{ where } r_{\text{CARO}} = [\text{kar}]$$

The two computed inflection zones  $\xi_{\text{masc}}$  and  $\xi_{\text{fem}}$  each produce two form realisation rules. The form realisation rules allow for building the four forms belonging to the inflecteme's paradigm.

$$\varphi_{\text{masc}}(\{\text{GENDER } \textit{masc}, \text{NUMBER } \textit{sg}\}) = r_{\text{CARO}} + m_{\text{masc-}sg} \text{ where } r_{\text{CARO}} + m_{\text{masc-}sg} = [\text{karo}]$$

$$\varphi_{\text{masc}}(\{\text{GENDER } \textit{masc}, \text{NUMBER } \textit{pl}\}) = r_{\text{CARO}} + m_{\text{masc-}pl} \text{ where } r_{\text{CARO}} + m_{\text{masc-}pl} = [\text{kari}]$$

$$\varphi_{\text{masc}}(\{\text{GENDER } \textit{fem}, \text{NUMBER } \textit{sg}\}) = r_{\text{CARO}} + m_{\text{fem-}sg} \text{ where } r_{\text{CARO}} + m_{\text{fem-}sg} = [\text{kara}]$$

$$\varphi_{\text{masc}}(\{\text{GENDER } \textit{fem}, \text{NUMBER } \textit{pl}\}) = r_{\text{CARO}} + m_{\text{fem-}pl} \text{ where } r_{\text{CARO}} + m_{\text{fem-}pl} = [\text{kare}]$$

Thus, the paradigm  $\mathfrak{P}_{\text{CARO}}$  of the Italian adjective CARO is:

$$\mathfrak{P}_{\text{CARO}} = \{([\text{karo}], \{\text{GENDER } \textit{masc}, \text{NUMBER } \textit{sg}\}), ([\text{kari}], \{\text{GENDER } \textit{masc}, \text{NUMBER } \textit{pl}\}), ([\text{kara}], \{\text{GENDER } \textit{fem}, \text{NUMBER } \textit{sg}\}), ([\text{kare}], \{\text{GENDER } \textit{fem}, \text{NUMBER } \textit{pl}\})\}$$

### References

- ANDERSON, Stephen R. (1992) *A-morphous Morphology*. Cambridge, UK: Cambridge University Press.
- ARONOFF, Mark (1994) *Morphology by Itself*. MIT Press.
- BAERMAN, Matthew (2006) Deponency in serbo-croatian. Online Database: <http://www.smg.surrey.ac.uk/deponency/Examples/Serbo-Croatian.htm>. Typological Database on Deponency. Surrey Morphology Group, CMC, University of Surrey.
- BAERMAN, Matthew (2007) "Morphological Typology of Deponency." In: M. Baerman/G. G. Corbett/D. Brown/A. Hippisley (eds.) *Deponency and Morphological Mismatches*, volume 145, p. 1-19. The British Academy Oxford University Press.
- BAERMAN, Matthew/CORBETT, Greville G. /BROWN Dunstan (2010) *Defective Paradigms*. Oxford, UK: Oxford University Press. Proceedings of the British Academy 145.
- BONAMI, Olivier/BOYÉ Gilles (2002) "Suppletion and dependency in inflectional morphology." In: F. V. Eynde/L. Hellan/D. Beerman (eds.) *The Proceedings of the HPSG '01 Conference*. Stanford, USA: CSLI Publications.
- BONAMI, Olivier/BOYÉ Gilles (2010) "La morphologie flexionnelle est-elle une fonction?" In: I. Choi-Jonin/M. Duval/O. Soutet (eds.) *Typologie et comparatisme. Hommages offerts a Alain Lemarechal*, p. 21-35. Louvain, Belgique: Peeters.

- BOYÉ, Gilles (2006) "Suppletion." In: K. Brown (ed.) *Encyclopedia of Language and Linguistics* (2nd ed.), volume 12, p. 297-299. Oxford, UK: Elsevier.
- BRESNAN, Joan (ed.) (1982) *The Mental Representation of Grammatical Relations*. MIT Press.
- CORBETT, Greville G. (2003) "Agreement: the range of the phenomenon and the principles of the Surrey database of agreement." *Transactions of the philological society* 101: 155-202.
- CORBETT, Greville G. (2007) "Canonical typology, suppletion and possible words." *Language* 83: 8-42.
- CORBETT, Greville G./FRASER Norman (1993) "Network Morphology: a DATR account of Russian nominal inflection." *Journal of Linguistics* 29: 113-142.
- FRADIN, Bernard (2003) *Nouvelles approches en morphologie*. Paris, France: Presses Universitaires de France.
- FRADIN, Bernard/KERLEROUX Françoise (2003) "Troubles with lexemes." In: G. Booij/ A. R. Janet de Cesaris/Sergio Scalise (eds.) *Selected papers from the Third Mediterranean Morphology Meeting, Topics in Morphology*, p. 177-196. Barcelona, Spain: IULA-Universitat Pompeu Fabra.
- KARTTUNEN, Lauri (1989) "Radical lexicalism." In: M. R. Baltin/A. S. Kroch (eds.) *Alternative Conceptions of Phrase Structure*, p. 43-65. Chicago: University of Chicago Press.
- MATTHEWS, Peter H. (1972) *Inflectional Morphology: a Theoretical Study Based on Aspect of Latin Verb Conjugation*. Cambridge, UK: Cambridge University Press.
- MATTHEWS, Peter H. (1974) *Morphology*. Cambridge, UK: Cambridge University Press.
- ROBINS Robert H. (1959) "In defense of WP." *Transactions of the Philological Society* 1959, p. 116-144.
- SAGOT, Benoît (2010) "The Lefff, a freely available, accurate and large-coverage lexicon for French." In: *Proceedings of the 7<sup>th</sup> Language Resource and Evaluation Conference*. Valletta, Malta.
- SAGOT, Benoît /WALTHER Géraldine (2011) "Non-canonical inflection : data, formalisation and complexity measures." In: *Proceedings of the workshop Systems and Frameworks in Computational Morphology (SFCM 2)*. Zurich, Switzerland.
- STUMP, Gregory T. (2001) *Inflectional Morphology. A Theory of Paradigm Structure*. Cambridge, UK: Cambridge University Press.
- THORNTON, Anna M. (2010) "Towards a typology of overabundance." Presented at the D écebrettes 7, Toulouse, France.
- WALTHER, Géraldine/SAGOT, Benoît (2011) "Modélisation et implémentation de phénomènes non-canoniques." *Revue Traitement Automatique des Langues* 52(2).
- ZAUNER, Alfonz (1973) *Praktická prirucka slovenského pravopisu*. Martin, Slovakia: Vydavate Istvo Osveta.
- ZWICKY, Arnold M. (1985) "How to describe inflection." In: M. Niepokuj et alii (eds.) *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistic Society*, p. 372-386. Berkeley, USA: Berkeley Linguistic Society.

Abstract  
MEASURING MORPHOLOGICAL CANONICITY

The question of regularity within morphological paradigms has been formerly addressed within approaches falling in the scope of *Canonical Typology* (Corbett 2003). The aim of this paper is to provide a means for assessing the notion of morphological canonicity through original measures developed within our new morphological framework  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ . In particular, we introduce original measures for non-canonical phenomena such as *heteroclisis*, *deponency*, *defectiveness* and *overabundance*.

We introduce  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  a new model for inflectional morphology using an inferential-realizational approach (Matthews 1974; Zwicky 1985; Anderson 1992). Our model precisely provides a formal representation of the lexicon/grammar interface. It relies on a formal definition of a lexical entry and a complete formal apparatus for computing all relevant form realisation rules for each lexeme, including stem formation rules. Realisation rules themselves may be expressed through any suitable realisation-based formalism (e.g. PFM or Network Morphology). We introduce several formal innovations such as *inflection zones*, that constitute partitions of given inflection classes. They are in particular used in modelling heteroclisis.

Povzetek  
MERJENJE MORFOLOŠKE KANONIČNOSTI

Vprašanja pravilnosti morfoloških paradigem so se že lotevali pristopi, ki sodijo v okvir *kanonične tipologije* (Corbett 2003). Cilj pričujočega članka je prispevati izvirne načine, ki bodo na podlagi meril, ki smo jih izdelali znotraj našega novega morfološkega modela  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ , omogočali ovrednotiti pojem morfološke kanoničnosti. Še posebej pozorno pa smo vpeljali nove načine merjenja nekanoničnih pojavov, kot smo npr. *heterokliza*, *deponentnost*, *nezapolnjenost*, *prenapolnjenost*.

V članku predstavljamo  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ , ki je nov oblikoslovni model, ki se opira na inferenčno-uresničitveni pristop (Matthews 1974; Zwicky 1985; Anderson 1992). Naš model ponuja prav formalno predstavitev slovarsko-slovničnega vmesnika. Temelji na formalni definiciji leksikalne iztočnice in popolnem formalnem aparatu, ki omogoča izpeljavo vseh relevantnih oblikoslovnih uresničitvenih pravil za vsak leksem, kamor sodijo tudi pravila oblikovanja osnove. Uresničitvena pravila lahko oblikujemo znotraj katerega koli ustreznega formalnega modela (na primer, teorija paradigmatških funkcij ali morfologija mrež /ang. *Network Morphology*/). Vpeljemo vrsto formalnih novosti, na primer *pregibna območja* (ang. *inflection zones*), ki tvorijo dele posameznega pregibnega razreda. Posebej koristni so pri modeliranju heteroklize.

## PHASES AT THE WORD LEVEL

### 1. INTRODUCTION

The central issue in this paper is how to account for the preservation of stress and vowel quality in English affixation and what differences exist between the »classic derivational« and Optimality Theory approach to the matter. The phenomenon in question is illustrated in (1). Words in English may but need not change the position of primary stress when affixed. In (1) the word *government* is derived from the word *govern* and the position of primary stress is not changed in derivation. When *governmental* is derived from *government*, primary stress shifts from the first to the pre-final syllable. In (2), primary stress is shifted in each of the two derived words, *instrumental* and *instrumentality*.

- |     |            |              |                 |
|-----|------------|--------------|-----------------|
| (1) | góvern     | góvernment   | gòvernmentál    |
| (2) | ínstrument | ínstrumentál | ínstrumentálity |

Furthermore, even if the position of primary stress is changed with affixation, the syllable bearing primary stress at some point in the derivation preserves the stress in the form of secondary stress (e.g. gòvernmentál, ínstrumentál) and is at the same time prevented from being reduced to a schwa (e.g. ìnstrumentálity).

In this paper we offer an analysis of the phenomenon in (1) that is a natural extension of the classic derivational analysis found in Chomsky and Halle (1968) (SPE hereafter). The latter is supplemented with a theory of word structure as in Marantz (2001) and will be referred to as a phase analysis of the preservation of stress, where the claim is that stress facts result from an interaction of stress rules, stress properties of affixes and word structure, and in which the preservation of stress in affixation is crucially connected with the word structure. The proposal together with the theoretical background it adopts will be presented in detail in sections 2-6. The paper will also present a conceptually very different analysis of the phenomenon, the Optimality Theory analysis of English word stress in Burzio (1994), comparing the two proposals in the relevant points, sections 7-8.

### 2. THEORETICAL BACKGROUND: THEORY OF WORD STRUCTURE

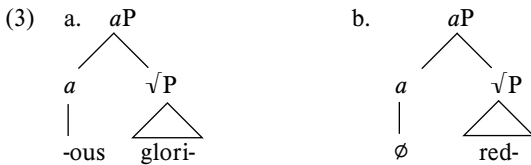
This paper is couched in the framework of Distributed Morphology, Halle and Marantz (1993), Halle (1997a), Marantz (1997), and its views on the word structure.

---

\* *Author's address:* Filozofska fakulteta, Oddelek za primerjalno in splošno jezikoslovje, Aškerčeva 2 Ljubljana, Slovenia. Email: tatjana.marvin@ff.uni-lj.si

The analysis in this study relies on the specific framework of word structure proposed in Marantz (1997) and Marantz (2001). In these two works Marantz argues for the unification of the inflectional and derivational morphology, which can be obtained if the two processes of word formation employ the same generative engine – the syntactic computation comprising the operations Merge, Move and Agree. In short, Marantz (2001) proposes that words are built by the Syntax performing all merger operations, including those between morphemes within a word.

A central innovation in Marantz (1997) is the treatment of roots and syntactic categories. In previous approaches to word formation syntactic categories such as V, N and A are properties of roots (stems) and affixes. In Marantz’s theory roots and affixes have no category per se, but are merged in the syntax with category-forming functional heads such as the ‘little’ *n*, *v*, *a* to form nouns, verbs and adjectives, respectively. These heads are typically realized by overt or phonologically null derivational affixes. For example, the adjectives *glorious* and *red* have the syntactic structures as in (3a,b), where  $\sqrt{\phantom{x}}$  stands for *Root* and the little *a* is filled with the derivational affix *-ous* in *glorious* and with a phonologically null affix in *red*.



### 3. PHASES AT THE WORD LEVEL

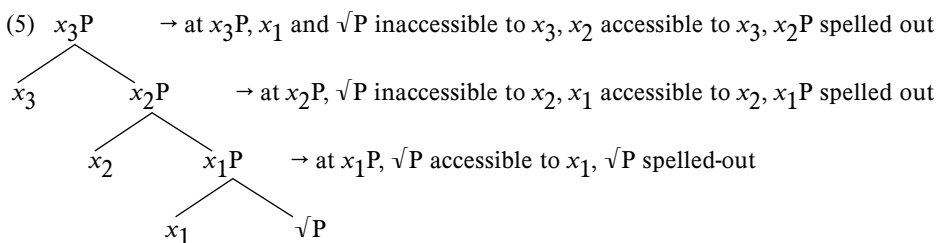
If words are treated as a result of syntactic processes, they are expected to show similar syntactic phenomena as found on the sentence level. One of such is Chomsky’s (2001) phase-by-phase spell-out, where the main idea is that derivation of sentences proceeds in phases, where phases can be seen as (predetermined) chunks of syntactic structure that are spelled-out at the phonological and semantic level and after that point impenetrable for potential semantic or phonological operations that the structure built on top of them might cause. Marantz (2001) proposes that phrases such as little *vP*, *nP* and *aP* constitute such chunks, i.e. phases at the word level. Category changing morphology realizing these little phrases within a word thus yields phases, and cyclic phonological effects within words such as the preservation of stress and vowel quality can be seen as following from the cyclic operation of syntactic derivation by phase, as in Marvin (2002). For example, if a syllable is assigned stress in a particular phase, this stress should somehow be preserved even if the word is affixed later on and new syllables carrying stress emerge. With this as a leading idea, we now investigate the interaction of stress assignment and the structure of words in English, arguing that stress and vowel quality in English are diagnostic of spell-out domains within words, and that in turn syntactic domains predict the spell-out. The analysis proposed is essentially a modernized version of the proposal found in SPE, where the notion of Chomsky and Halle’s »cycle« corresponds to the »phase

at the word level« and the »phase spell-out« relates to the SPE notion »preservation of stress from earlier cycles«.

The idea that category-forming phrases constitute phases can be summarized as follows:

- (4) a. Phrases headed by word-forming functional heads, such as little *v*, little *n* and little *a*, constitute spell-out domains at the word level, Marantz (2001).
- b. Phrases at the word level are subject to Chomsky’s (2001) Phase Impenetrability Condition.
- c. *Phase Impenetrability Condition at the word level*: H and its edge (specifiers, adjoined elements) are spelled out at the next strong phase. The domain of H is spelled out at the phase of HP. A head *h* adjoined to H is in the domain of H.

At the point of the merger of the category forming head *x* (where *x* stands for *v*, *n*, or *a*, and is supplied with derivational affixes), the complement of the little *x* is spelled out (i.e. meaning and pronunciation are determined) and from that point on inaccessible to heads attaching higher. The idea is schematized in (5).<sup>1</sup>



#### 4. STRESS ASSIGNMENT IN ENGLISH: FROM CYCLES TO PHASES

The standard example from SPE illustrating the phenomenon of the preservation of stress and vowel quality in English affixation is the ‘minimal pair’ *condensation - compensation*. SPE observes that in some dialects of English the boldfaced /e/ in *condensation* reduces to a schwa, while this is not the case in the word *compensation*, despite the fact that phonotactically and morphologically the two words are very similar. The vowel reduction is a consequence of the Vowel Reduction Rule, which reduces a lax vowel to a schwa in English. The explanation for this fact offered in SPE is that the difference between the two nominalizations follows from the stress of their constituents. That is, the nominalizations ‘contain’ verbs *condense* and *compensate*, which represent the first cycle in the formation of the nouns *condensation* and *compensation* and which have different stress patterns.<sup>2</sup>

<sup>1</sup> See Embick (2010) for a slightly different execution of the spell-out mechanism.

<sup>2</sup> The SPE notion ‘cycle’ can be roughly defined as the point of the application of stress rules to a certain constituent.

- |     |                         |                      |   |
|-----|-------------------------|----------------------|---|
| (6) | a. <i>condens</i> ation | (/e/ is not reduced) | <i>cond</i> ense → <i>cond</i> ensation   |
|     | b. <i>compens</i> ation | (/e/ is reduced)     | <i>c</i> ompensate → <i>compens</i> ation |

In the former verb, primary stress is found on *condense*, which consequently means that the stressed vowel cannot be reduced to a schwa in the cycle of the verb, where stress is assigned for the first time. In the latter verb, main stress is found on the first syllable of *compensate*, therefore the corresponding /e/ in *compensate* can be reduced to a schwa in the cycle of the verb. If the stress from earlier cycles is preserved, we now have a natural explanation for the different vowel quality in the two nominalizations. In *condensation*, the vowel /e/ has received stress on an earlier cycle, i.e. the verb cycle, and is therefore prevented from being reduced. In *compensation*, the vowel /e/ has been reduced on the earlier cycle and has remained the same in the nominalization.<sup>3</sup>

This paper links the insight from SPE to a finer grained syntax at the word level as proposed in Marantz (2001), producing what I term the phase analysis. In the phase analysis the phase spell-out and Phase Impenetrability Condition as in (4) apply to stress in connection with vowel reduction. Stress and vowel quality preservation in the sense of SPE and Kiparsky (1979) are a consequence of the phase spell-out mechanism—when a chunk of word to which the stress rules have applied is spelled out, the new affixation and Vowel Reduction cannot erase this information.

The stress in English words is a result of the interaction of the three factors in (7).

- (7)
- a. Building blocks with selectional and stress-related specification
  - b. Set of phonological rules
  - c. Structure of building blocks

The building blocks (affixes) have stress related specification, which means that they can be stress neutral or can trigger the application of stress rules when attached. The set of phonological rules comprises different rules of English stress and the order in which they apply (see section 4.1. for a detailed discussion). Some phenomena, such as the preservation of stress in derivation, cannot be accounted for only by the properties of building blocks and stress rules, requiring for their explanation also the third element, the structure of the building blocks. Specifically, we will try to show that when stress has been assigned in a particular chunk (little *x* in our analysis), neither the properties of building blocks attached to the chunk nor the stress rules can erase this information – once assigned, stress is preserved.

#### 4.1. Preview: English stress in general

English is a language in which word stress depends on phonological properties and the position of syllables. Descriptively, English words fall into three groups with respect to stress, SPE, Burzio (1994), Halle (1998). Nouns and simple adjectives typ-

---

<sup>3</sup> As to the Vowel Reduction Rule, SPE places it after the process of stress assignment within the word, i.e. in the non-cyclic block.



ically follow pattern (8a), while verbs and derived adjectives typically follow pattern (8b). There is a third group that consists of words with one primary and one secondary stress, as in (8c).<sup>4</sup>

- (8) a. Heavy penultimate: *agénda*, *appéndix*, *horizon*  
 Antepenultimate: *américa*, *ásterisk*
- b. Superheavy final: *prevént*, *decíde*, *obéy*, *annóy*  
 Penultimate: *inhábit*, *imáginé*, *paréntal*, *governméntal*
- c. *málachite*, *eléctróde*, *monophýsite*, *cóllòid*

The analysis of these three groups that I shall adopt is as in Halle (1998). Relying on the metrical theory in Idsardi (1992) and Halle and Idsardi (1995), Halle (1998) proposes that the English stress system is constituted by the Main Stress Rule (henceforth MSR) supplemented by two edge-marking rules. The MSR has two parts.<sup>5</sup> The first part constructs a binary foot at the end of a string whose last asterisk projects a light syllable. If the last syllable is heavy or there are not enough syllables in the word to construct a binary foot, a unary foot is constructed. This part of the MSR is illustrated in (9). For example in the word *develop*, the last syllable is light, therefore a binary foot is constructed: *de(velop*. In the words *usurp* and *cajole*, on the other hand, the last syllables are heavy, therefore a unary foot is constructed: *u(surp*, *ca(jole*. In the rightmost column we find words with only one syllable, where only unary feet can be constructed regardless of the syllable weight: (*put*, (*black*.

(9)	*(**	*(*	*(*	(*	
	<i>devélop</i>	<i>usúrp</i>	<i>cajóle</i>	<i>pút</i>	
	<i>clandéstine</i>	<i>robúst</i>	<i>divíne</i>	<i>bláck</i>	Halle (1998)

In addition to being subject to the MSR, the words illustrated in (10) and (11) are also subject to an edge-marking rule (henceforth EMR), which applies to a list of words before the MSR. The first of the two edge-marking rules, referred to as the *RLR* EMR, inserts a right parenthesis before the final syllable of the word if the syllable contains a short vowel, (10).

<sup>4</sup> The secondary stress in derived words is not due to membership in group (8c).

<sup>5</sup> Here I adopt a theory of stress following Liberman (1975), Prince (1983), Halle and Vergnaud (1987), Idsardi (1992), Halle and Idsardi (1995) and Halle (1997), in which stress contours of words are expressed by means of a metrical grid. The bottom line (line 0) is composed of projections of the stressable elements (syllables), and higher lines constructed by projecting certain of these elements upward (the heads). Feet are constructed by boundaries (left or right parentheses); they are left or right headed.

(10)	*(**)*	*(*)*	*(*)*	(*)*	
	América	agéndá	Tacóma	vílla	
	cómpetent	consistent	cohérent	cúrrént	
	oríginal	paréntal	anecdótal	móral	Halle (1998)

The second edge-marking rule, the *LLR* EMR, inserts a left parenthesis to the left of the rightmost syllable and applies in words where *RLR* EMR has not applied. This is illustrated in (11).<sup>6</sup>

(11)	(**[*	*(**[*	**(**[*	([**	
	málachíte	stalágmíte	monophýsite	Hússite	
	amýgdalòid	mollúscòid	epicýclòid	cóllòid	
	plálatinòid	eléctróde		ánòde	Halle (1998)

The EMR rules summarized in (12) apply first to a list of words; they are followed by the the MSR in (13),<sup>7</sup> which applies to all words. Note that the feet constructed on line 0 are left-headed.

(12) **Edge-marking rules**

a. RLR Edge Marking

$\emptyset \rightarrow ]$  in env. \* \_\_ \* ## line 0

Condition J: Final asterisk projects short vowel.

b. LLR Edge Marking

$\emptyset \rightarrow [$  in env. \* \_\_ \* ## line 0

(13) **Main Stress Rule**

a.  $\emptyset \rightarrow ($  in env. \_\_ \* \* <P\* > line 0

Condition K: Second asterisk projects vowel in a light rime.

b.  $\emptyset \rightarrow ($  in env. \_\_ \* <P\* > ## line 0

Halle (1998)

<sup>6</sup> Examples in (11) have two stresses, one primary and one secondary. They are subject to the Rhythm Rule, Halle (1998), which places the main stress on the leftmost syllable on line 1 of the metrical grid. The Rhythm Rule in Halle (1998) is an extended Rhythm Rule proposed by Liberman and Prince (1977), applying in word sequences as well as within single words.

<sup>7</sup> A few other minor rules are needed to account for certain groups of exceptions, for which the reader is referred to Halle (1998). Also, it is important to note that the rules proposed by Halle (1998) have a few lexical exceptions, as exemplified in (i) and (ii). These words are marked in the lexicon as special and cause problems to any theory of English stress.

- (i) (\*)\* also: *modest, solemn, modern, auburn, covert* (unexpectedly subject to RLR)  
góvern
- (ii) (\*\*)\*  
Cáthol-ic also: *Arabic, politic* (unexpectedly subject to RLR)

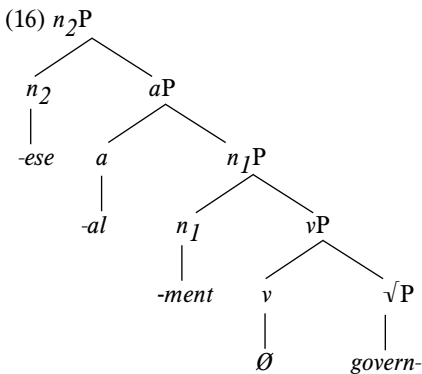
## 5. RULE BASED APPROACH AND PHASES APPLIED TO ENGLISH STRESS

In this part, we finally join the three elements summarized in (7) and claim the following:

- (14) a. Words are composed of little xPs, Marantz (2001);  
 b. MSR and EMR apply at every xP if triggered by a diacritic marking on *x*;  
 c. Vowel Reduction Rule takes place at the level of ‘prosodic word’;  
 d. A ‘phase analysis’ (phase spell-out and Phase Impenetrability Condition) as in (4) applies to stress in connection with vowel reduction.

To illustrate the application of (14), the word *governmentalese* in (15) is fully derived in (17a-e). Following Marantz (2001), the structure of this word is as in (16).

- (15) *gòvernmentàlése*  
meaning: complicated or obscurantist language thought to be characteristic of government bureaucratic statements  
pronunciation: /gʌvərnmentəli:z/



At each phase the corresponding metrical grid is indicated after the illustration of rule application by using a diacritic, (e.g. *á*), to mark stressable elements that receive asterisks on line 1. The Vowel Reduction rule (which is non-cyclic in nature) takes place after the spell-out of  $n_2P$ .

At  $vP$ , where the root phrase is spelled out, the stress rules apply, which gives the root spell-out *govern*, illustrated in (17a).

- (17) a. at  $vP$ : *govern-(∅)* → EMR → MSR → the root is spelled out as *gòvern*

line 1 \*  
 line 0 (\* ]\*  
           gòvern

At the next phase,  $n_1P$ , the affix *-ment* is added to the structure and consequently to the metrical grid, but since this affix is non-cyclic, it does not trigger application of EMR and MSR. Thus the  $vP$  is spelled out as *govern*, as illustrated in (17b).

(17) b. at  $n_1P$ : *govern(ment)* → the  $vP$  is spelled out as *govern*

```

line 1      *
line 0      (* * *
            govern ment

```

The next spell-out, illustrated in (17c), occurs at  $aP$ , where the stress rules are triggered by the attachment of the cyclic affix *-al*. The underlying representation of the affix *-ment* is still accessible at this point, so after that syllable receives an asterisk on line 1, the vowel in *-ment* surfaces as a / $\epsilon$ /. The underlying representation of the root is inaccessible at this point, so even if the stress rule at this point assigns no asterisk on line 1 to the root *govern-*, the latter still retains the asterisk due to its already having been spelled out two phases ago, as in (17a). The resulting effect of this mechanism is that it appears as if the stress is preserved from the previous assignments.

(17) c. at  $aP$ : *government(al)* → EMR → MSR → the  $n_1P$  is spelled out as *gouvernmént*.

```

line 1      *      *
line 0      (* * (* ] *
            govern ment al

```

At  $n_2P$ , the complement  $aP$  is spelled out as *gouvernméntal*, as illustrated in (17d). The affix *-ese* attached at this point triggers a reapplication of the stress rules, causing a projection of a line 1 asterisk on the syllable /*ese*/. However, since the chunk *gouvernmént-* is inaccessible to the stress rules applying at  $n_2P$  all line 1 asterisks of  $n_1P$  are preserved in the grid.

(17) d. at  $n_2P$ : *governmental(ese)* → EMR → MSR → the  $aP$  is spelled out as *gouvernméntal*

```

line 1      *      *      *
line 0      (* * (* * (*
            govern ment al ese

```

Finally, at the next higher phase  $n_2P$  is spelled out. Stress rules have assigned a line 1 asterisk to *-ese*, while line 1 asterisks of  $aP$  are still preserved from previous phases. Since now we reach the end of the word, Vowel Reduction takes place. This is illustrated in (17e).

- (17) e. at the next higher phase: *governmentalese*: → EMR → MSR → Vowel Reduction  
 →  $n_2P$  is spelled out as *góvernmentálése* /gʌvərnməntəli:z/

If the asterisks assigned by stress rules on previous phases were erased, the Vowel Reduction rule would reduce the vowels in /gov-/ and /ment-/. However, these chunks of the word were spelled out in previous phases with the vowels in question receiving line 1 asterisks. Given Phase Impenetrability Condition, this information is automatically carried over to the last phase and thus these vowels are prevented from the Vowel Reduction.<sup>8</sup>

## 6. SOME CONCEPTUAL ISSUES IN PHASE ANALYSIS

### 6.1. Advantages

An analysis that makes use of a phase spell-out on the word level certainly has some advantages over theories that do not employ the word structure in the derivation of stress facts. First, in a phase analysis the stress preservation insights from SPE and Kiparsky (1979) can be directly linked to a finer-grained syntactic approach as in Marantz (2001), where the structure of words plays an active role in predicting the stress in derived words. A more global property of this approach is also the fact that a single generative engine is employed in producing words, phrases and sentences in contrast to the so-called lexicalist approaches that utilize two separate generative systems, one for words and one for phrases and sentences, e.g. DiSciullo and Williams (1987).

### 6.2. Problematic issues

As always, several problematic issues need to be taken into account. The first one is that not all stress-related processes can be subject to Phase Impenetrability Condition as in (4), one of such being the assignment of primary stress, which may change with affixation (e.g. *góvernment* - *gòvernmentál*), even though stress is preserved once assigned – the first syllable of the word *governmental* carries a secondary stress as a result of once having been assigned a primary stress on that syllable in the word *government*. Similarly, the so called Rhythm Rule, Liberman and Prince (1977), Kiparsky (1979), can change the properties of the already spelled-out constituent, as can be seen in (18) for the word *thirteen*. Given these facts, it has to be the case that primary stress assignment is a phenomenon occurring at lines higher than line 1 in the metrical grid and is therefore not subject to the phase spell-out and the Phase Impenetrability Condition.

- (18) a. thirtéen  
 b. thirtèen mén

---

<sup>8</sup> In this derivation, we do not state explicitly what the difference between primary stress and subsidiary stress is, the term stress covering both of the two specific cases. See section 6.2. for a more detailed discussion in connection with this relationship.

The next issue to be discussed is that if one adopts Halle (1998) rules of stress assignment, one still depends on the idiosyncratic property of the individual affix as to whether its phase will or will not trigger the application of stress rules. For example, *-ment* as a stress neutral-affix does not trigger the stress rules, while *-al* as a stress-changing affix does. It would be perhaps desirable if every phase also meant the application of stress rules, however, this cannot be achieved in the present theory.

The last issue to be discussed is of empirical nature and relates to a small group of exceptions that cannot be explained by a phase analysis. If in example (19) the right-hand word is derived from the left-hand one, where the latter is already an xP, then we run into a problem of derived words either losing or acquiring stress on the parts that should be already spelled out.<sup>9</sup>

- |   |   |
|---|---|
| (19) a. <i>sólid</i> – <i>solídity</i>    | b. <i>télegràph</i> – <i>tèlégraphy</i>     |
| c. <i>catástrophe</i> – <i>càstróphic</i> | d. <i>compónent</i> – <i>còmponentiáliy</i> |

For example, in the word *solidity*, consisting of the adjective *solid* and *-ity*, we expect the stress to be as in *sólidity* with no vowel reduction of /ò/, but instead we get *solídity* with a reduced vowel. This problem is acknowledged already in SPE, where a solution is offered in the form of a series of rules applying to specific groups of words such as *solidity*. The important thing to note is that the groups of words that behave as *solidity* are not the general case, but exceptions that are very hard to capture in any theoretical background. In this paper I rely on the SPE treatment of exceptions as being somehow marked in the lexicon as exceptions to the standard set of rules that otherwise apply to a vast majority of words.<sup>10</sup>

## 7. ENGLISH STRESS IN OPTIMALITY THEORY

In this part we present the basic overview of a non-rule based approach, with special emphasis on stress assignment in derived words. The analysis of English stress considered is Burzio (1994), which belongs to the set of Optimality Theory-based analyses. Burzio’s account makes use of interactive well-formedness constraints in the spirit of McCarthy and Prince (1993) and Prince and Smolensky (1993), where stress is not assigned by phonological rules, but is rather a result of the ranking of constraints which impose conflicting demands on representations. This section is intended as an illustration of the main principles rather than a detailed summary of the whole of Burzio’s (1994) proposal.

---

<sup>9</sup> The words that acquire subsidiary stress in the derivation of the type *condéense-còndénsátion* and the words that lose subsidiary stress of the type *consúlt-cònsúltátion* can be accounted for by relating their stress to their internal structure. See SPE and Marvin (2002) for a proposal.

<sup>10</sup> “There are many details and special cases that do not seem to fall under any large-scale generalizations and that shed little light on general questions of phonological theory or on the structure of English.”, (SPE: 113).

## 7.1. English word stress in general

To begin with, let us summarize Burzio's (1994) analysis for underived words. Burzio proposes a typology of possible feet as in (20), where »H/L« stands for heavy/light syllables and vowels followed by colons are long.

(20) Possible feet	Non-rightmost	Rightmost
a. (Hσ)	mo(nòn ga)hé:la	àri(zó: na)
b. (σLσ)	(winnepes)sáukee	a(mé ri ca)

Burzio (1994)

The second component of his analysis is English specific existence of a special class of syllables, referred to as weak, which may or may not be metrified, as exemplified in (21), where the weak syllables are italicized. In the right-hand column, the final weak syllable is extrametrical, which is why these words appear as exceptional in terms of stress.

(21) a. a ris (to cra cy) / (ac cu ra)cy	
b. ob(ject <i>tive</i> ) / (ad ject) <i>ive</i>	Burzio (1994)

Weak syllables are special also when not extrametrical. They form the so-called weak feet, which fail to attract primary stress, as illustrated in (22). In English, primary stress falls on the rightmost non-weak foot.

(22) a. (órtho)(dòxy)	
b. (árchì)(tècture)	Burzio (1994)

The final postulate of the analysis is that all English words end in a vowel, leading to positing final null vowels in words such as *robust* or *develop*, as shown in (23).

(23) a. ro(búst∅) / (éarnes)t∅, (fré quen)ce	
b. de(vé lo p∅) / (ás te ris)k∅	Burzio (1994)

## 7.2. Constraint ranking

In this part the main constraints that regulate stress assignment in English are presented, first for non-affixed and then for affixed words.

### 7.2.1. No affixation

The highest ranked constraint is the so-called Metrical Well-formedness, which represents a set of possible feet in English and is summarized in (24):

(24) Metrical Well-formedness, Burzio (1994: 165):

	Feet	Non-rightmost	Rightmost
a.	mo(nòn.ga)héla	(Hσ)	(Hσ) a(gén.da)
b.	(wìn.ne.pes)sáukee	(σLσ)	(σLσ) a(mé.ri.ca)
c.	ac(cé.le)râte	(Lσ)	#(Lσ) h(ónes)t

Indeterminacies with respect to parsing, e.g. whether a certain chunk is parsed as (σLσ) or σ(Lσ) or to the metrification of weak syllables are resolved by further constraints that are ranked lower than Metrical Well-formedness. These are summarized in (25) as part of the Metrical Alignment constraint, Burzio (1994: 166) and will not be further summarized in this paper.

(25) *Metrical Alignment*

- a. Strong Retraction condition: ... (σσ)(HW)#
- b. Metrical Alignment: \*(σ...H...)
- c. Metrification of verbs: ...∅#
- d. Exhaustive Parse: #(.../...)#

### 7.2.2. *Affixation*

Burzio (1994) notes the two familiar patterns of stress preservation under affixation, which he refers to as weak stress preservation and strong stress preservation, exemplified in (26) and (27), respectively. The first one is traditionally argued to occur when stress-changing suffixes are attached, while the second occurs with stress-neutral suffixes.

(26) napóleon na(pòle)ónic

(27) pròpagánda pròpa(gándis)t∅

When it comes to the interaction between stress and word-formation, Burzio (1994) argues that a major player in this domain is the principle that imposes consistent metrical characteristics on morphemes and in preservation of stem stress under affixation. The constraint regulating the preservation of stem stress under affixation is Metrical Consistency, which is aligned as in (28) with respect to the two other stress assigning constraints in his system.

(28) Metrical Well-formedness >> Metrical Consistency >> Metrical Alignment

In this system, Metrical Well-formedness is the most important property and no English suffix can evade metrification in any way. The difference between what we traditionally term stress-neutral and stress-changing suffixes (i.e. between (27) and (26)) is a reflex of the way in which the indeterminacy of final weak syllables can combine with the phonological structure of the suffix itself. Stem stresses are pre-



served if and only if they correspond to independently well-formed feet. In the word *napoleonic*, which contains a stress-changing affix *-ic*, the stem *napoleon* and the affix *-ic* interact in such a way that the constraint of Metrical Well-formedness requires a remetrification such as in (29).<sup>11</sup>

- (29) ...ic  
 a. ...σ0 (L1 L2 σ3) → ...σ0 (L1 L2)(σ3 i cə)  
 b. na(póle)(ónicə)

In the word *propagandist*, with a stress-neutral affix *-ist* and the stem *propaganda*, the syllable from the affix overlaps with the syllable of the stem and the constraint of Metrical Well-formedness assigns stress as in (27) in a predictable fashion; no remetrification is needed, which results in an apparently complete preservation of stress.<sup>12</sup>

In Burzio's system the default case is that stress does not change with affixation unless the phonological properties of the stem and affix combination require a remetrification of the whole to obey the higher ranked well-formedness constraint. Affixes are thus neither stress neutral nor stress changing by nature as in the SPE approach, but appear as such due to the interaction with the stem and the tendency to obey the well-formedness constraint.

## 8. A comparison to a phase analysis

In this part we compare Burzio's analysis with the phase analysis as proposed in this paper, with special emphasis on the stress preservation phenomenon. To repeat, Burzio (1994) deals with stress assignment in English by combining metrical well-formedness and stem stress preservation in an Optimality Theory fashion (constraints in (28)), while the phase analysis as in Marvin (2002) utilizes a rule-based approach towards English stress as in Halle (1998), where stress preservation is connected to the word structure through the notion of word level phase, Marantz (2001). We shall see that although Burzio's analysis deals elegantly with the notion of stress neutral and stress changing affixes, eliminating this property by making it a result of stress related constraint ranking, the phase analysis crucially captures the connection between structure of words and stress when it comes to affixes with a mixed behavior as to causing a stress change – these are affixes that sometimes appear as stress-changing and sometimes as stress-neutral. In such cases Burzio's analysis fails to see the structure-stress connection, which in his system comes out just as a by-product of constraint ranking and affix-stem combination interplay.

---

<sup>11</sup> Other cases of weak preservation of stress are dealt with in a similar fashion in Chapter 6-7 of Burzio (1994). For the reasons of space we cannot give further examples in this paper.

<sup>12</sup> Other cases of strong preservation of stress are dealt with in a similar fashion in Chapter 8-9 of Burzio (1994). For the reasons of space we cannot give further examples in this paper.

## 8.1. Stress preservation and mixed suffixes

When it comes to stress preservation, Burzio (1994) introduces a constraint on stress preservation, which makes sure stress is preserved in affixation, but which can be superseded by the Metrical Well-formedness constraint, the result of which is what appears to be a stress shift in some cases. For Burzio such stress shift has traditionally been mistakenly ascribed to the properties of the affixes (whether they are stress changing or stress neutral). In the phase analysis, the phenomenon of cyclic preservation of stress is explained as a result of the stress rules, the properties of individual affixes and the structure. As one can see, albeit indirectly, Burzio (1994) appeals to word structure, introducing a constraint by which affixed words preserve stem stress, though all work in stress placement is done by the metrical consistency constraint and the connection between the constraint and the word structure appears random. It should also be noted that Burzio (1994) does not see stress preservation linked to vowel reduction in contrasting pairs such *cond[e]nsation* and *comp[ə]nsation*. He claims that the difference in vowel length follows from constraints of the preservation of segmental quality (from *cond[e]nse* and *comp[ə]nsate*) that is independent of preservation of stress. In this paper I limit the comparison between his analysis and the phase analysis to the position that the two take to the preservation of stress.<sup>13</sup>

Let us now turn to the case of mixed suffixes, where the two approaches to stress preservation crucially differ in their treatment. Mixed suffixes, such as *-able* or *-ist*, are suffixes that in some words appear as stress changing (e.g. *rémedy-remédiable*), while in other words they appear as stress neutral (e.g. *prevént-prevéntable*).

Burzio (1994) solves this puzzle by referring to the metrical properties of the combination of the stem and the affix, claiming that *-able* is stress-neutral if the first syllable of *-able* replaces the null vowel, as in (30).

- (30) a. *pre(vént)∅* → *pre(vénta)ble*  
b. *in(hábit)∅* → *in(hábita)ble*

Otherwise, *-able* appears stress-changing, since it requires remetrification as in (31).

- (31) a. *(démon)(strá:te)* → *de(mónstra)ble*                      b. *(ímpreg)(nàte)* → *im(prégna)ble*  
c. *(rémedy)* → *re(mé:dia)ble*                                      d. *(súrvey)* → *sur(véya)ble*

Burzio (1994) also notes the following generalization: mixed suffixes are systematically neutral with free stems but not neutral with bound ones, the property which in his system is captured by positing the two constraints as in (32).

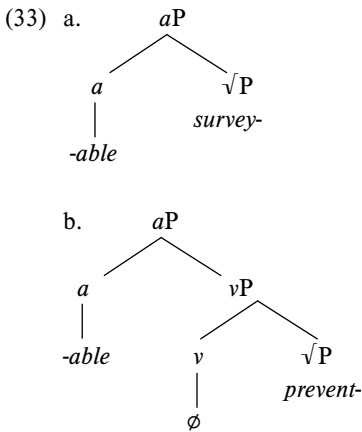
---

<sup>13</sup> For a detailed proposal, the reader is referred to chapters 4 and 10 in Burzio (1994) and to Burzio (2007). For problematic cases, such as the pair *inf[ɔ]rm – inf[ə]rmation*, which are elegantly dealt with by referring to the word structure in SPE and Marvin (2002), Burzio has to introduce further constraints to handle the non-preservation of segmental quality.

- (32) Metrical consistency hierarchy: a > b  
 a. Stem consistency  
 b. Suffix consistency
- Burzio (1994)

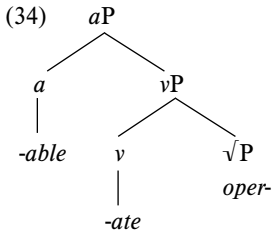
Mixed suffixes appear neutral when attached to a free stem, satisfying (32a) (e.g. *americanist* derived from *american*), but impose a fixed stress pattern otherwise, i.e. with bound stems, (e.g. *antagonist* derived from *antagon-*), satisfying (32b), when (32a) is irrelevant.

In the phase analysis the »mixed« behavior of suffixes such as *-able* is directly related to the structure of the word; it is a consequence of the attachment site of the affix. Namely, *-able* appears stress-changing in (33a) because it is introduced in the first phase (i.e. with bound stems); otherwise, when introduced on top of another phase (i.e. with free stems), it appears stress-neutral, (33b):



Similarly, such analysis explains the fact that base verbs in *-ate* are often (e.g. *comparable*), but not always truncated, as in *cultivable*, *emancipatable*, *operatable*, Plag (2003). The stress pattern of the non-truncated adjectives is predictable (*òperátable*) and follows from the structure in (34). The affix *-able* is not attached to the root directly, but to an already made verb *òperàte*, or in other words, it is introduced on top of the first phase, *vP*, and has thus, given (4), no bearing on the stress assigned in the previous phase.<sup>14</sup>

<sup>14</sup> The primary stress assignment is not subject to (4), thus the changes in the position of primary change are allowed. What is important for this analysis is that stress is preserved, albeit as subsidiary stress.



The fact that mixed affixes appear stress-neutral with free stems and stress-changing with bound ones does not come about coincidentally in the phase analysis. This fact is seen as a cause rather than a consequence—the affix *-able* is in principle always stress-changing, however, it can only bring about the stress change if attached in a certain structural position (e.g. it can change the root stress is attached to the root directly, but not otherwise, (34 or 33b)).

### 8.2. Stress preservation in multiple suffixation

A phase analysis also deals straightforwardly with the following generalization from Burzio (1994), where he observes that neutral suffixes maintain their word integrity under further suffixation—if in the structure (35) *suf<sub>1</sub>* is neutral, then *suf<sub>2</sub>* can affect at most *suf<sub>1</sub>*, but not the stem, as seen from the example (36).

- (35) stem + *suf<sub>1</sub>* + *suf<sub>2</sub>*
- (36) *desire* → *desirable* → *desirability*

In the phase analysis such generalization is directly predicted. Given (4) and (5), attached material can influence the stress of the previous phase, but not of the phase below the previous phase. In (35) it is therefore expected that *suf<sub>2</sub>* could influence the stress placement of the material in the *suf<sub>1</sub>*, but not of the stem, because the latter will be spelled out at the point where *suf<sub>1</sub>* is attached and thus become inaccessible for the changes that further suffixation might bring about. Again, what follows directly from the structural position of the affix in the phase theory is coincidental in the Optimality Theory analysis.

## 9. CONCLUSION

In this paper we presented an analysis of English stress preservation in derived words that is a natural extension of the classic derivational analysis found in SPE, supplemented with a theory of word structure as in Marantz (2001). The main idea of the proposal is that stress facts in English result from an interaction of stress rules, stress properties of affixes and word structure, and where the preservation of stress in affixation is crucially connected with the word structure. The paper also briefly presented an Optimality Theory analysis of English word stress, Burzio (1994), com-

paring the two proposals in how they deal with the preservation of stress phenomenon. To deal with the latter, Burzio (1994) introduces a constraint on stress preservation, which makes sure stress is preserved in affixation, but which can be superseded by the Metrical Well-formedness constraint, the result of which is what appears to be a stress shift in some cases. For Burzio such stress shift has traditionally been mistakenly ascribed to the properties of the affixes (their being either stress changing or stress neutral) and this is one of the points where the two proposals differ crucially. Another crucial difference between the phase and the Optimality Theory analysis is the predictions that the two make as to the so-called mixed suffixes. The phase analysis can account straightforwardly for the mixed behavior as following from the structural position that the affix takes in the word, while in the Optimality Theory analysis such instances appear either coincidental or are made to follow from further constraints that indirectly rely on the structure of words.

## References

- BURZIO, Luigi (1994) *Principles of English Stress*. Cambridge, UK: Cambridge University Press.
- BURZIO, Luigi (2007) "Phonology and Phonetics of English Stress and Vowel Reduction." *Language Sciences* 29, 154-176.
- CHOMSKY, Noam (2001) "Derivation by phase." In: M. Kenstowicz (ed) 1-52.
- CHOMSKY, Noam/Morris HALLE (1968) *The Sound Pattern of English*. New York: Harper and Row.
- DISCIULLO, Anna-Maria/Edwin WILLIAMS (1987) *On the Definition of Word*. Cambridge, MA: MIT Press.
- EMBICK, David (2010) *Localism versus Globalism in Morphology and Phonology*. Cambridge, MA: MIT Press.
- FUDGE, Eric (1984) *English Word-Stress*. London: George Allen and Unwin.
- HALLE, Morris (1997a) "Distributed Morphology: Impoverishment and Fission." *MIT Working Papers in Linguistics* 30, 425-449.
- HALLE, Morris (1997b) "On stress and accent in Indo-European." *Language* 73, 275-313.
- HALLE, Morris (1998) "The Stress of English Words 1968-1998." *Linguistics Inquiry* 29, 539-568.
- HALLE, Morris/Jean-Roger VERGNAUD (1987) *An Essay on Stress*. Cambridge MA: MIT Press.
- HALLE, Morris/Michael KENSTOWICZ (1991) "The Free Element Condition and cyclic versus noncyclic stress." *Linguistic Inquiry* 22, 457-501.
- HALLE, Morris/Alec MARANTZ (1993) "Distributed Morphology and the Pieces of Inflection." In: K. Hale/S. J. Keyser (eds), 111-176.
- HALLE, Morris/William IDSARDI (1995) "General Properties of Stress and Metrical Structure." In: J. Goldsmith (ed), 402-443.
- HALLE, Morris/Alec MARANTZ (1994) "Some Key Features of Distributed Morphology." *MIT Working Papers in Linguistics* 21, 275-288.
- HAYES, Bruce (1985) *A Metrical Theory of Stress Rules*. New York: Garland Publishing.

- IDSARDI, William (1992) *The Computation of Prosody*. Doctoral dissertation. MIT.
- KIPARSKY, Paul (1979) "Metrical Structure Assignment is Cyclic." *Linguistic Inquiry* 10, 421-441.
- KIPARSKY, Paul (1982) "From Cyclic Phonology to Lexical Phonology." In H. van der Hulst/N. Smith (eds), 130-175.
- LIBERMAN, Mark (1975) *The Intonational System of English*. Doctoral dissertation. MIT.
- LIBERMAN, Mark/Alan PRINCE (1977) "On Stress and Linguistic Rhythm." *Linguistic Inquiry* 8, 249-336.
- MARANTZ, Alec (1997) "No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon." In: A. Dimitriadis/L. Siegel/C. Surek-Clark/A. Williams (eds), 201-225.
- MARANTZ, Alec (2001) "Words." ms., MIT.
- MARVIN, Tatjana (2002) *Topics in the Stress and Syntax of Words*. Cambridge, MA: MIT Working Papers in Linguistics.
- MARVIN, Tatjana (2005) "English syllabification and schwa-insertion: From the *Sound pattern of English* to the notion of phase." Ljubljana: *Linguistica*, 97-106.
- MCCARTHY John/Alan PRINCE (1993) "Prosodic Morphology, I: Constraint Interaction and Satisfaction." ms., University of Massachusetts, Amherst and Rutgers University, New Brunswick.
- PLAG, Ingo (1999) *Morphological Productivity*. Berlin: Mouton de Gruyter.
- PRINCE, Alan (1983) "Relating to the Grid." *Linguistic Inquiry* 14, 19-100.
- PRINCE, Alan/Paul SMOLENSKY (1993) "Optimality Theory: Constraint Interaction in Generative Grammar." ms. Rutgers University, New Brunswick, and University of Colorado, Boulder.

#### Abstract

#### PHASES AT THE WORD LEVEL

This paper presents a view on word structure in which words are argued to exhibit the syntactic phenomenon of Chomsky's (2001) phase-by-phase spell-out in Marantz (2001) fashion, where category-forming phrases such as little *vP*, *nP* and *aP* constitute phases at the word level. Category changing morphology within a word thus yields phases, and cyclic phonological effects within words can be seen as following from the cyclic operation of syntactic derivation by phase, as in Marvin (2002). The paper investigates the interaction of stress assignment and the structure of words in English, arguing that stress and vowel quality in English are diagnostic of spell-out domains within words, and that in turn syntactic domains predict the spell-out. The analysis proposed is essentially a modernized version of the proposal found in Chomsky and Halle (1968), where the notion of Chomsky and Halle's "cycle" corresponds to the "phase at the word level" and the "phase spell-out" relates to the "preservation of stress from earlier cycles". The proposal is then compared to an Optimality Theory analysis of English word stress found in Burzio (1994), especially with respect to the analysis of stress preservation in English word-formation. Burzio (1994) deals with stress assignment in English by combining metrical well-formedness and stem stress preservation in the form of interacting constraints. In his system, contrary to the one in the phase analysis, the relationship between word stress and word structure appears random.

Povzetek  
FAZE NA BESEDNEM NIVOJU

Članek obravnava pojav ohranitve naglasa v angleških izpeljankah v luči morfološke teorije, ki predpostavlja, da so besede zgrajene po postopkih skladenjske komponente slovnice in so kot take podvržene t.i. faznemu izpisu ("phase spell-out"), kot ga za stavčni nivo predlaga Chomsky (2001), za besedni nivo pa Marantz (2001). V tem pristopu so besede pojmovane kot besedne zveze t.i. malih glagolskih, samostalniških in pridevniških jeder *g*, *s* in *p*, ki jih tipično udejanjajo glagolske, samostalniške in pridevniške pripone pri izpeljavi. Male besedne zveze so faze, kar pomeni, da na stopnji, ko jih kot sestavnike dodamo v drevesno strukturo besede, hkrati izpišemo njihovo dopolnilo, ki je po tem trenutku nedostopno za fonološke in pomenske spremembe od nadaljnjih sestavnikov, ki se priključijo besedi. V takem pristopu lahko ciklične fonološke pojave, kot je na primer ohranitev besednega naglasa pri izpeljavi v angleških besedah, izpeljemo s pomočjo skladenjske strukture. Analiza s fazami na besednem nivoju je posodobljena analiza, ki jo za ciklične pojave najdemo v Chomsky in Halle (1968). V članku jo nadalje primerjamo z analizo angleškega besednega naglasa v okviru optimalnostne teorije, ki jo najdemo v Burzio (1994). Burzio (1994) predlaga, da je naglas v angleščini posledica omejitev glede možnih površinskih oblik, povezava med naglasom in besedno zgradbo pa je v njegovi teoriji le stranski produkt medsebojnega vpliva omejitev.





## FROM VERBAL PREFIXES TO DIRECTION/RESULT MARKERS IN ROMANCE

### 1. VERB - LOCATIVE CONSTRUCTIONS IN ROMANCE<sup>1</sup>

It is well known that standard Romance languages, following the classical Latin tradition, employ prefixes – more or less opaque – and make a restricted use of verb-locative (or verb-particle) constructions, in order to express spatial information with verbs.<sup>2</sup> Italian, French, Spanish, Romanian, Catalan and Portuguese grammars and dictionaries register prefixed verbs, and only marginally the corresponding analytic constructions, with a verb followed by a locative, forming one lexical element.<sup>3</sup> Differently, analytic constructions (called in English “phrasal verbs”, in German “trennbare Verben”) are very common in Germanic languages.<sup>4</sup>

However, the evolution of the Latin syntactical type SOV into the present SVO type – which was fundamental to the weakening of the case system and has led to the formation of prepositions to express syntactical relations – has affected the prefix system too, making it less transparent and less productive.<sup>5</sup> Thus, when we look to diatopic<sup>6</sup>, diachronic<sup>7</sup>, diaphasic and diastratic<sup>8</sup> varieties of Romance, we note that they frequently substitute prefixed verbs with verbs plus locative constructions. This phenomenon has been discussed mainly on Italian data, since in colloquial Italian it is largely attested and various kinds of verbs followed by locatives are employed<sup>9</sup>.

---

\* *Author's address:* Facoltà di Lettere e Filosofia, Dipartimento di Studi Letterari Linguistici e Filologici, Via S.Croce, 65, Trento, Italia. Email: patrizia.cordin@unitn.it

<sup>1</sup> I use the term ‘locative’ both for referring to places and to directions.

<sup>2</sup> Verb-locative constructions are attested in non classical Latin texts; see Durante (1981) and Vincent (1999).

<sup>3</sup> On standard Italian, Spanish and Catalan grammars and dictionaries see Calvo (2010).

<sup>4</sup> A recent overview on the construction in Germanic languages is presented in Dehé et alii (2002).

<sup>5</sup> On the decline of the Latin system of verbal prefixation and the progressive loss of transparency and productivity see Iacobini (2004) and Masini (2006: 81-84).

<sup>6</sup> On northern Italian dialects and Ladin see Rohlfs (1969), Gsell (1982), Spiess (1983); more recently the construction in northern Italian varieties has been discussed by Vicario (1997), Benincà/Poletto (2006), Vigolo (2007), Cini (2008), Cordin (2011). Acadian French verbal constructions with locatives are given in King (2000).

<sup>7</sup> Old French examples are discussed in Dufresne et alii (2003); for ancient Italian see Masini (2006) and Jezek/Strik-Lievers (2010); for old Italian dialects see Cordin (2011).

<sup>8</sup> See Iacobini/Masini (2007).

<sup>9</sup> Among others, see Schwarze (1985), Simone (1997), Jansen (2004), Cini (2008).

In fact, in spoken Italian we find:

- a) stative verbs + locative: *stare su* = *raddrizzarsi* ‘to stand up’, *stare via* = *assentarsi* ‘to be away’;
- b) intransitive motion verbs + locative:<sup>10</sup> *andar fuori* = *uscire* ‘to go out’, *andar su* = *salire* ‘to go up’, *correr dietro* = *rincorrere* ‘to chase’, *venir dentro* = *entrare* ‘to come in’;
- c) transitive motion verbs + locative: *buttar via* = *eliminare* ‘to throw away’, *metter dentro* = *inserire* ‘to put into’, *tirar su* = *raccogliere* ‘to pick up’;
- d) figurative motion verbs + locative: *tirarsi su* = *diventare allegri* ‘to get better’, *metter sotto* = *soggiogare* ‘to keep somebody down’ or *investire* ‘to knock somebody down’; *tirar avanti* = *sopravvivere* ‘to survive’, *andar fuori* = *ammattire* ‘to be out of one’s mind’;
- e) activity verbs + locative: *grattar via* = *raschiare* ‘to scrape away’, *lavar via* = *ripulire* ‘to wash away’, *strappar via* ‘to rip off’, *cancellar via* = *scancellare* ‘to delete’.

The verbs in e), which in Romance area seem peculiar to Italian, deserve a particular attention: the locative *via* ‘away’ is used in combination with verbs of manner that contain a semantic feature of removing. The adverb expresses not only the direction that is involved as a component of the activity, but – as Iacobini/Masini (2007) notice – it gives an aspectual information, suggesting a telic reading, that emphasizes the success of the process and that underlines its result, changing an activity into an accomplishment.

The aspectual function of the adverb in verb-locative constructions is more evident in some northern Italian dialects, where – as we will see in next paragraph – it is not limited to verbs that semantically encode a feature of removing and is not restricted to the adverb *via* ‘away’.

## 2. FROM SPACE TO ASPECT

The verb-locative construction is largely attested in some dialects of northern Italy (in particular Friulan, Veneto, Piedmontese, Lombard, Trentino), where locative adverbs appear not only in combination with motion verbs (excluded manner motion verbs), but also with different kinds of transitive verbs and unaccusative verbs, where locatives contribute to give some aspectual interpretations of the event. In these dialects locatives combine frequently with transitive verbs where an action is performed on an affected object: their interpretation is still connected (although indirectly) to directionality, in that they specify the end-point of a movement that characterizes the action, usually carried out with the hands or arms, as already noticed for Italian verbs like *raschiare via*. In dialectal varieties, however, the ‘resultativization’ through the combination of a verb and a locative is much more frequent, as the following examples taken from Trentino illustrate.<sup>11</sup> I group them under e’):

---

<sup>10</sup> Manner motion verbs, however, cannot co-occur with a locative, as it will be illustrated in footnote 22.

<sup>11</sup> More examples – taken from ALTR (2005) and Grassi (2009) – are discussed in Cordin (2011).

e') *coertàr su* to cover with a blanket; to cover something, eg wood, with a canvas etc., to shelter it, *copiar ò* to copy out, *desfàr giu* to wind up, *destèndar fò* to unfold, to spread out, *destrigàr fora* or *destrigàr via* to tidy up, put things away, make space, *far su* to build, to put up, to raise, *far ò* to slice (potatoes, apples), *fregàr su* to polish the floor, *inciavar via* to lock away, *lavàr fò* to wash out a large container, *lavàr giu* to do the washing up, *lavàr su* to clean the floor, *misiàr su* to stir, *netàr fò* to do the clearing, *rasàr via* to rub (away), *rebaltàr su* to turn, to knock over, *refàr su* to rebuild, *regolàr via* to tidy (away), *schiciàr*, *scuciàr* o *strucàr fò* to squeeze something hard to get the liquid that it contains out, to squeeze, *scónder via* to cover up a thing, so that it can't be seen, *scrivar giu* to write down, *sgrafàr ò* to scrape off, *spazàr su* to sweep, *spazàr ò* to brush down an item of clothing, *sugàr giu i piàc*, *le pošade*, *i bicér* to dry up (after the washing up), *sugàr su* to dry a wet table or floor, *tajàr giu* to cut, to chop off; to slice, to cut in slices, *zoncàr via* to cut, to chop off

The telic interpretation is more evident in those constructions where locatives (mainly *su* 'up' and *fora* 'out', 'off') combine with unaccusative verbs or with other transitive verbs that do not express actionality carried out physically with hands and directed to a specific endpoint, as in e). The telic constructions given in f) often require as direct object a plural noun, or a singular collective noun, or a quantified noun (particularly the pronoun *tutto*, *tutti* 'all', or the adjective *tutto* + noun):

f) *brontolàr su* to grumble, *comeðàr via* to agree, to keep something quiet, *comeðarse fora* to work out, to come out right (of a situation), *contàr fora tut* to blurt up, to blab, to gossip, *contàr su* to tell, *contentar via* to keep happy, to please, *dir fora* to blab, *dir su* to talk big, to tell stories, to spin a good yarn, *emblogiarse su*, *engranizarse su* to dirty, to soil, to smear oneself, eg with soot, *finìr fò* to get through (everything), to use up, *fracasàr su tut* to break everything, *imbotonarse su* to button up (one's cloths), *magnàr fò* to polish off (a meal, food), *morìr fora* to die (back, off), *morìr ò* (of fire) to go down, to die down, *pagàr fora* to liquidate, to pay off, to sell out, *scaldar fora* to heat, *studiar fora* to examine, *spozzarse su* to become black, to darken, to get filthy, *squèrzirse su polito* to wrap up well, to dress warmly, *stropàr su tut* to shut up a house before leaving it, *vènder fora tut* to sell out, *vestirse su polito* to wrap up well.

The frequency with which the quantifier *tutto* appears bears out the hypothesis that telicity can be linked with the idea of measuring an event, and with a scale or path which provides a means to determine the progress of an event.

Another aspectual meaning – closely connected to the measurement of an event – is expressed by the adverb *su* in conjunction with an action verb. The adverb suggests a certain speed in the execution of an action, leading to a loss of accuracy and a consequently unsatisfactory result. Therefore the expressed aspect is always telic, linked to a measurement of the event, with the goal of an action only partially achieved, as shown below:

g) *co me còntet su?* ‘What are you saying? (I find it hard to believe that)’, *far su* ‘to be up to no good’, *farse su* ‘to dress in bad taste’, *giustàr su* ‘to patch up’, *rangiàr su a la mèio* ‘to adjust the least worst’, *slambrotàr su* ‘to speak unclearly’, *todescàr su* ‘to speak German to someone who doesn’t understand’

The adverbs *su* and *fora* can also indicate the repetition of an action, or its duration. In both cases there is a measurement of the event, as illustrated by the constructions *basarse su* ‘to kiss repeatedly’, *struforlarse su* ‘to be all over each other’, *brazarse su* ‘to embrace each other’, in which the verbs describe physical contact. Also constructions with verbs related to speech, thought, sight, eg *pensarghe su* ‘to reflect, to think about’, *dir su la corona* ‘to say the rosary’, *contar su na storia* ‘to tell a story’, *vardàr fò i so mistér* ‘to mind one’s own business’ indicate intensity or meticulousness.

In a small (but frequently used) group of Trentino constructions the locative adverb expresses the progress of an event rather than its result: ingressive and progressive readings in particular are attested. A locative expresses an ingressive aspect when it indicates the beginning of an action; *via* ‘away’, *su* ‘up’, ‘on’, often accompany and reinforce verbs which themselves express the idea of beginning: *ciapàr su* ‘to be on the point of, to be about to’, *inviàr via* ‘to start up, to begin’, *mòver via* ‘to start something off, to open (a restaurant, etc.)’, *scomenzàr via* ‘to start off’, *stradàr via* ‘to start off’. The progressive aspect is expressed by locatives which indicate that an action is in progress, and is rendered solely by the adverb *dietro* ‘behind’ in combination with the verb *esser* ‘to be’ + *a* + infinitive, or *esser* + *che* ‘that’ + finite verb: ‘to be busy doing something’:

h) *co sét dré a far?* ‘what are you doing?’; *l é dré che l cubia* ‘he’s sleeping’; *ié dré che i và* ‘they’re leaving’; *ié dré che i và en malga* ‘they’re going up to the summer pastures’; *l’é dré che la sbèrcis* ‘she’s dying’; *la vaca l’é dré a far* ‘the cow is about to calve’; *le róse l’é dré che le sbòcia* ‘the roses are blooming’; *l é dré ala lègna* (lit. he is behind) ‘he is working a) to collect wood b) to stack wood’; *l ò gatà che l éra dré a caminàr* ‘I found him while he was leaving’

The variety of Trentino examples given from e) to h) shows that this dialect – along with others of northern Italy – has fully developed the possibilities of interpretation for the locative in the construction verb-locative, attesting a systematic passage from prefixed verbs to verbs plus locative, both with directional and with aspectual interpretation.

This tendency – although present in all Romance languages that have no rigid codification – might have been strengthened by the requirement of northern Italian dialects for spelling out inflection markers. In particular, Trentino seems to require

also the spelling out of a goal/result marker that functions both for arguments of motion verbs and for aspect.

We know that in Italian an argument can be left unexpressed, if the context (discursive or situational) allows it to be understood. Verbs of motion often contain an unexpressed argument, usually relating to either the endpoint or the origin of their path:

- 1a            *Entra pure* (qui, in questa stanza)!  
Come in (here, into this room)!
- 1b            *Quando esco* (di qui, da questa casa), *chiudo sempre la porta*.  
When I go out (of here, of this house), I always shut the door.
- 1c            *È partito* (dalla città) *alle due in punto*.  
He went away (from the city) at two o'clock.

We can hypothesize that, unlike in Italian, in Trentino dialects the position of the arguments of verbs of motion must always be marked by an explicit element (which is spelled out), a locative, in fact – even in cases in which the argument noun phrase is not expressed. The spelled out element acts to mark the position of a verb of motion's internal argument. The Italian examples given in 1a-c are in fact correct in Trentino dialect, with a locative which signals the presence of the argument internal to the verbal phrase (goal, or origin), as shown in 2a-c:

- 2a            *Vèi dentro, dai!*  
Come in!
- 2b            *Quando che vago fora, sero sempre la porta*  
When I go out, I always shut the door.
- 2c            *L'è nà via ale doi precise*.  
He went away at two o'clock.

Some evidence of the requirement for marking the position of a direction argument in Trentino is given by those examples in which the locative combines with a non prefixed verb whose meaning is characterized by an inherent feature of natural directionality, as *cascar zo* 'to fall down', *levarse su* 'to stand up', *scampar via* 'to run away', *sentarse zo* 'to sit down', *montar su* 'to go up': it is clear that the locative is required not to express directionality, which is already inherent to the verb, but rather to mark the position of the argument of the motion verb.

Further confirmation of the "argument marker" hypothesis comes from the cases in which the locative governs an unexpressed noun, whose understanding, although not provided contextually, is still made possible thanks to the verbal semantic features: *méter su* 'to put on (something on the fire, to cook)', *tor for ua* 'to take grapes, choosing them from the vineyard', *tor fora (i cavai)* 'to unhitch (horses from a cart etc.)', *tor fora (i soldi)* 'to withdraw (money)', *deśniàr fò* 'to flush,

drive out an animal'. Here, too, the locative acts as the marker of an argument position (goal or origin), whose content is left implicit.<sup>12</sup>

Trentino dialects seem to require a marker not only as a signal an argument of motion verbs, but also as a marker of a dynamic feature of an event, as it happens in particular with telic verbs<sup>13</sup>. To place these observations in a wider context, I note that the requirement of an explicit marker for directional arguments and for verbal aspect can be seen as a particular instance of a rule requiring that specifiers of functional verbal categories (agreement, aspect, argument structure, modality) are spelled out. Frequently the spelling out of these elements originates doubling. As a consequence, in northern Italian dialects also the verb-locative combination can be dealt with coherently within the framework of doubling, which is relevant in these diatopic varieties, as illustrated by the Trentino examples 3-7:

Subject agreement doubling<sup>14</sup>

3 *el Mario el zuga*

the Mario 3<sup>rd</sup> p.sing. plays

Mario plays

Dative clitic doubling<sup>15</sup>

4 *ghe dago la torta ai popi*

I dat.clit.3<sup>rd</sup> p.pl. give the cake to the kids

---

<sup>12</sup> Yet another confirmation of the locative as argument-marker comes in Tortora's 2002 study of the Borgomanero dialect in Piedmont. In this dialect the locative behaves differently in sentences with object clitics in which it functions as an argument of a transitive verb of movement, and in sentences in which, on the other hand, it functions as the circumstantial complement of a transitive verb. When it is an argument, the locative occupies the position immediately to the right of the verb, before the object clitic; when it is an adjunct it occupies the position immediately to the right of the object clitic. The following examples with the locative *denti/denta* (inside), taken from Tortora (2002: 726) illustrate the above distinction:

i\* *i porta-la denti*

ii *i porti denta-la*

I'm bringing it inside

iii *i moengia-la denti*

iv\**i moengia denta -la*

I'm eating it inside

Note that Borgomanerese shows general clitic enclisis with both finite and infinite verbs.

<sup>13</sup> Tortora (2002) suggests that clauses may contain an aspectual phrases field, where multiple aspectual phrases instantiate different components of aspect and that the locatives in the constructions under discussion can be placed in the aspectual phrases field of a clause. Tortora thus develops a hypothesis originally presented by Cinque (1999) about those Italian adverbs like *già* 'already' and *più* 'anymore' that contribute to the terminative aspectual interpretation and the perfective aspectual interpretation of the event. See also Poletto (2000) about the existence of an agreement field in the clause.

<sup>14</sup> See Brandi/Cordin (1989).

<sup>15</sup> See Cordin (1993).

I give the cake to the kids

Accusative clitic doubling with pronouns

5 *el m'ha ciamada mi*

he me-clitic has called me

he called me

Negation doubling

6 *No vegno miga con ti*

I don't come *miga* (neg.) with you

I'm not coming with you

Interrogative modality doubling

7 *Nen-te o sten-te?*

Should we go-te (inter. clitic) or stay-te (inter.clitic)?

Should we go or stay?

### 3. COMPARING ROMANCE AND GERMANIC VERB - LOCATIVE CONSTRUCTIONS

The Trentino constructions quoted in §2 present many analogies with German *trennbare Verben*, as it is illustrated in table 1, where a comparison is given between the two mentioned languages, Italian and English.

Table 1

Trentino	Standard Italian	German	English
nar via	uscire	ausgehen	to go out
molar zo	abbassare	herunterlassen	to let down
tirarse su	raddrizzarsi	aufrichten	to stand up
montar su	risalire	einsteigen	to go up
far su	fare, rifare	umbauen	to build, to put up
tiràrghe dré (a qualchedun)	assomigliare	nach jemandem schlagen	to take after someone
vènder fora	svendere	ausverkaufen	to sell out
morir fora	estinguere	aussterben	to die (off)
scorlar zo	scuotere	abschütteln	to shake out
giustar su	aggiustare	ausbessern	to fix
enviar via	avviare	einleiten	to start (off)
taiar via	ritagliare	abschneiden	to cut off

The analogy between German *trennbare Verben* and the dialectal verb-locative constructions employed in northern Italy has been noticed by various authors, who have

proposed that the Romance construction is the result of the influence of the German language on those Romance varieties which are more exposed to the linguistic contact.<sup>16</sup> However, as we have already seen, the construction is not a quirk of northern Italian dialects, but the consequence of a tendency attested in the whole Romance area -although with different degrees of productivity- and this suggests that the Romance construction may be considered as an independent process. The hypothesis is also confirmed by ancient Romance languages, where this type of construction was widespread.<sup>17</sup>

Verb-locative combinations in Germanic and in Romance languages have also different morphological and syntactic properties: it has been noted that only Germanic languages permit nominal derivations from the construction, whereby Romance languages do not permit them.<sup>18</sup> As for the syntactic properties, we know that in many Germanic languages (German, Norwegian, Danish, Icelandic) the inflected verb of a main clause moves from last into second position, leaving the locative alone. As a consequence, the two elements are usually separated by one or more phrases. By contrast, in Romance constructions the locative comes immediately after the verb.<sup>19</sup> Romance locatives can only be separated from the verb by clitic objects or “light” adverbs (temporal/aspectual: *ancora* ‘still’, *mai* ‘never’, *sempre* ‘always’, *mica* ‘not at all’).<sup>20</sup>

---

<sup>16</sup> See Plangg (1980) and Gsell (1982). A different account is presented by Kramer (1987).

<sup>17</sup> See Dufresne et alii (2003); Masini (2006); Jezek/Strik-Lievers (2010); Mateu/Rigau (2010); Cordin (2011).

<sup>18</sup> For instance, in Italian *la sottomissione dei ragazzi* ‘the subduing of the children’ derives from the phrase *sottomettere i ragazzi*, ‘to subdue the children’; however, from the phrase *mettere sotto i ragazzi* it is impossible to derive *la missione (messa) sotto dei ragazzi*. Similarly, in Northern Italian varieties we find: *ridere* ‘to laugh’ from which *risada* ‘laugh (the noun)’ derives, and *ridere drio* ‘to deride’ from which one cannot derive *\*risada drio* ‘derision’.

<sup>19</sup> The adjacency of locatives to the verb could be accounted for considering them as elements belonging to a category intermediate between (+clitics) and (-clitics). In fact in the verb-locative constructions locatives, although tonic, share various properties with clitics: they cannot appear outside their clause (topicalized), they cannot be modified, and they cannot be coordinated with other elements. Following a proposal made in Cardinaletti (1995) and in Cardinaletti/Starke (1999) for some German and Italian pronouns and for some adverbs, we could call locatives combining with verbs “weak adverbs”.

<sup>20</sup> Consider the examples v and vi, respectively with a clitic and with a light adverb before the locative:

v Non mandarlo fuori.  
Do not send him out.

vi Tira sempre/ ancora /già fuori la lingua.  
He sticks always /again / already out his tongue.

Neither direct and indirect objects (vii, viii), nor “heavy” adverbs (ix) can appear between the verb and the locative:

vii\* Hai messo una moneta dentro.



Finally, as a further confirmation of the independence of the Romance periphrases from German, I recall the Trentino examples given in the group h), §2 that are formed with *esser + dré/drio*, ‘to be after/behind ...’, and have a progressive reading: they are not found in German varieties, therefore it is impossible to recognize them as calques.

Recently, from a different point of view, some scholars, comparing Germanic verb-locative constructions and Romance ones, have re-examined the typological distinction proposed by Talmy (1991, 2000) between verb framed languages and satellite languages.<sup>21</sup> According to Talmy’s generalization, Romance languages should be considered verb framed, in that they express motion using verbs, whereas Germanic languages should be considered satellite languages, in that they express motion using a satellite around the verb, e.g. a particle. Iacobini and Masini (2007) argue against this generalization claiming that Italian – at the light of the data quoted in §1 – “does not conform to Talmy’s generalization, since it behaves more like English than Spanish.” Mateu/Rigau (2010: 242) reach a different conclusion, proposing that “Italian *verbi sintagmatici* resemble Germanic phrasal verbs, but only superficially”. According to Mateu/Rigau, the crucial point is that in Romance non directional manner verbs cannot appear with a particle expressing path, while in Germanic languages pure manner verbs can co-appear with the locative particle, as illustrated by the contrast between the examples a-b in 8 and 9:

- 8a. John danced the night off.  
 b\* \*Gianni ha danzato fuori tutta la notte.
- 9a. John worked his debt off.  
 b \*Gianni ha pagato fuori i suoi debiti.

In order to account for this difference, Mateu/Rigau (2010: 253) posit that Romance constructions obey a restriction, being allowed only where a concrete or abstract “path-feature” is spelled out by two different roots, both in the verb and in a directional particle. It follows therefore that “the list of verbs that enter into the Romance verb-particle constructions is reduced to those verbs encoding or involving

---

viii\*    Darebbe a suo fratello contro.  
 \*He would give his brother against.

ix \*    Lo hai messo come niente fosse sotto.  
 \* You have knocked him without noticing down.

<sup>21</sup> Among others, see Folli/Ramchand (2005); Horrocks/Stavrou (2007); Iacobini/Masini (2007); Spreafico (2009); Mateu Rigau (2010).

directionality.”<sup>22</sup> Differently, Germanic particle constructions are also allowed with pure manner verbs.

This hypothesis, however, does not account immediately for those Trentino verb-locative constructions quoted in §2 (groups f, g), where locatives combine with verbs that do not encode directionality, or at least that do not encode only directionality.<sup>23</sup>

In order to explain all Romance cases, it seems in fact necessary to specify what the feature “directionality” involved in the verb means. We see that in many examples of northern Italian varieties the interpretation of directionality is not only rela-

---

<sup>22</sup> Folli/Ramchand (2005) point out that in Italian only motion verbs involving a result feature can express the particle. They compare path verbs (*correre* ‘to run’, *rotolare* ‘to roll’, *rimbalzare* ‘to bounce’, *scivolare* ‘to slide’, *gattonare* ‘to crawl’, *saltare* ‘to jump’, *volare* ‘to fly’, *saltellare* ‘to hop’) that involve the direction feature and allow the verb-particle construction (*correre fuori*, *rotolare giù*, *rimbalzare su*, *scivolare via*, *gattonare via*, *saltare su*, *giù*, *fuori*, *volare via*, *saltellare su*) with manner of motion verbs (*galleggiare* ‘to float’, *camminare* ‘to walk’, *galoppare* ‘to gallop’, *danzare* ‘to dance’, *nuotare* ‘to swim’, *sciare* ‘to ski’, *passeggiare* ‘to walk around’, *vagabondare* ‘to wander’) that do not involve directionality and do not allow the construction. I note that Romance motion verbs that do not allow verb-particle constructions are unergative (in Italian they require the auxiliary *avere* ‘to have’), in contrast with those motion verbs that allow the construction and are unaccusative (in Italian and in Trentino they require the auxiliary *essere* ‘to be’). The contrast is clearly illustrated by the following couple of examples, x and xi:

x Gianni è corso su al parco.

John is run (has run) up to the park.

xi \*Gianni ha corso su al parco tutto il giorno.

John has run up to the park the whole day long.

We see that the same verb has a different behavior with the auxiliary *avere* ‘to have’ and with the auxiliary *essere* ‘to be’: in fact, only the unaccusative *correre* with the auxiliary *essere* allows a verb-locative construction.

<sup>23</sup> One of the examples that Mateu/Rigau (2010: 243) quote for illustrating the contrast between the Germanic and the Romance construction is 8a, that I repeat here:

8a John danced the night off.

Like Italian, Trentino disallows this construction. Nonetheless, it is interesting to note that in this variety we can find a very similar construction (8c), where the adverb *fora fora* literally ‘off off’ appears, with an aspectual meaning corresponding to “in a continuous way, without interruption”.

8c El Gianì l’à balà fora e fora tuta la not.

John has gone on dancing all the night.

*Fora fora* in Trentino also appears in combination with other activity and motion verbs, as *pianzer* ‘to cry’, *nodar* ‘to swim’, ecc. (*l’ha pianzest fora e fora per tut el viazo* ‘he went on crying during the whole journey’, *l’ha nodà fora e fora sin a la costa* ‘he went on swimming until he reached the coast’). The aspectual interpretation of the adverb *fora fora* is registered in Aneggi 1984, who translates the expression into the Italian *di continuo, sempre* ‘continuously, always’. Analogously, another locative adverb with doubling *drio a drio* ‘behind behind’ is used in Trentino for expressing verbal aspect, as registered by Groff 1955, who translates it as ‘di seguito, ininterrottamente’ (consecutively, step by step).

tive to space, but it is also relative to aspect: in fact it gives a sort of measure of the development of a process, expressing the endpoint (the success in reaching a complete result, or a partial result), or the start-point, or the intensity of the activity, or its progress. Only extending the interpretation of directionality in order to cover both space and time (phases of time), it is possible to give a coherent account for Romance verb-locative constructions, included the quoted Trentino examples, maintaining Talmy's distinction, according to Mateu/Rigau (2010).

#### 4. CONCLUDING REMARKS

I summarize here the most relevant points discussed in paragraphs 1-3. First of all, I consider verb-locative construction a pan-Romance phenomenon. Even though the prefixal system typical of classical Latin has been codified in standard Romance languages, where it is still productive, non standard varieties attest a wide and varied employ of the verb-locative construction.

Apparently this construction is analogous to the Germanic verb-particle one, but a comparison between the two languages shows that the Romance one is the result of an independent process and is characterized by different morphological and syntactical properties: unlike Germanic, Romance languages do not allow nominal derivations from the construction; moreover, only in Romance constructions the locative must be adjacent to the verb (or come immediately after a clitic object, or after a "light" adverb) and cannot be separated from the verb by any constituent.

Further evidence comes from the diachronic analysis of Romance documents which confirm that the phenomenon in Romance area appears independently from the influence of Germanic languages, being attested since the beginning of vulgar languages and being not limited to calques.

Even though the linguistic contact between German and Romance varieties cannot be considered the cause of the diffusion of the construction in Romance area, it might have contributed to reinforce an autonomous tendency. In fact, the construction is more largely attested in the Rhaeto-Romance area and in the northern Italian regions that have German neighbours (Trentino, Friuli, the Ladin Dolomites), as well as in Veneto, in Lombardy and on the Appenines around Parma, which are regions that have had historical ties with Austria.

The comparison between different Romance varieties, and in particular between spoken Italian and northern Italian dialects, evidences a significant cross-linguistic variation in Romance area. The construction can be ranked along a continuous scale of directionality, whose poles are clearly defined: at one end we find the most transparent constructions with compositional meanings, which are formed by a verb of motion plus a locative expressing location or, more often, direction in space; at the other pole, we find those combinations that have non-compositional meanings, and are formed by a verb followed by a locative that has lost its spatial meaning, and has assumed a grammatical one, expressing an aspectual meaning. The intermediate verbs (group e, e') alternate between a concrete directional reading and a grammaticalized one. Only in some varieties, as in Trentino dialects (more in general in northern

Italian varieties), the construction has reached the grammaticalization pole, where the locative functions as an explicit marker of a grammatical feature of the verb.<sup>24</sup>

Those varieties in which the construction has reached the most grammaticalized pole are characterized by the tendency of spelling out inflection elements, such as the agreement with the subject, with the dative object, etc. As a consequence, they frequently show doubling effects (subject doubling, case doubling, negation doubling).

Doubling leads to constructions containing an element which seems to be semantically superfluous. In verb-locative combinations, too, the locative often repeats a semantic feature inherent to the verb that seems redundant. Redundancy is often symptomatic of semantic erosion, and it is the first sign of a system's loss of productivity. The weakening of a number of systems (of subject agreement,<sup>25</sup> morphological case, prefixes as expressions of aspect) may be the link between locative constructions and the doubling phenomena: it seems possible in fact that Trentino dialects make the functional features of a verb explicit in a new position in order to strengthen a weak system.

Further confirmation to the possible linking between doubling phenomena and verb-locative constructions in Romance comes from the recent discussion of the typological distinction between Germanic languages, in which the expression of events of motion is considered satellite-framed, and Romance languages, in which it is considered verb-framed. According to Mateu/Rigau (2010: 257), the verb-particle construction with path motion verbs in Romance languages can be accounted for by a structure in which the direction particle specifies the abstract path that has been copied into the verb: the same syntactic feature can be spelled out twice, both on the verbal root and on a "prepositional-like element".

Summarizing, Romance languages have maintained the possibility typical of Latin verb prefixes of expressing both space and aspect directionality in the verb. However, in their evolution some varieties that have been less codified have developed a new system of spelling out this feature, using a locative adverb after the verb in order to mark a direction/result feature, that is either the argument position of a motion verb or the verbal aspect of an event.

---

<sup>24</sup> I propose the following schema:

	v-loc=spatial direction	v-loc=end-point of the action	v.loc=measure of the result
French, Spanish	+	-	-
Italian	+	+	-
Trentino	+	+	+

<sup>25</sup> Consider that in Trentino dialects verb endings for the III person singular are the same as those for the III person plural.

## References

- ALTR - CORDIN, Patrizia (ed) (2005) *L'Archivio lessicale dei dialetti trentini*. Trento: Dipartimento di Scienze Filologiche e Storiche.
- ANEGGI, Aldo (1984) *Dizionario cembrano (triangolo Sover-Montesover-Piscine): parole e cose - frasi - modo di dire - proverbi del dialetto della valle di Cembra*. S. Michele all'Adige (Trento): Museo degli usi e costumi della gente trentina.
- BENINCÀ, Paola/POLETTI, Cecilia (2006) "Phrasal Verbs in Italian and Regional Italian." In: F. Hinskens (ed) *Language Variation. European Perspectives*, Selected Papers from the third International Conference on Language Variation in Europe, Amsterdam, June 2005. Amsterdam: John Benjamin, 9-22.
- BOOIJ, Geert/VAN KEMENADE, Ans (2003) "Preverbs: an introduction." In: G. Booij/J. van Marle (eds) *Yearbook of Morphology*. Dordrecht: Kluwer, 1-11.
- BRANDI, Luciana/CORDIN, Patrizia (1989) "Two Italian Dialects and the Null Subject Parameter." In: O. Jaeggli/K. Safir (eds) *The Null Subject Parameter*. Dordrecht: Kluwer Academic, 111-142.
- CALVO, Cesareo Rigual (2010) "I verbi sintagmatici italiani, con appunti contrastivi con lo spagnolo e il catalano." In: M. Iliescu/H. Siller Runggaldier/P. Danler (eds), vol. 7, 375-384.
- CARDINALETTI, Anna/STARKE, Michael (1999) "The typology of structural deficiency: a case study of the three classes of pronouns." In: Riemsdijk H. van (ed) *Clitics in the languages of Europe*. Berlin, de Gruyter, 145-233.
- CARDINALETTI, Anna/MUNARO, Nicola (eds) (2009) *Italiano, italiani regionali e dialetti*. Milano: FrancoAngeli.
- CINI, Monica (ed) (2008) *I verbi sintagmatici in italiano e nelle varietà dialettali. Stato dell'arte e prospettive di ricerca*. Atti delle giornate di studi (Torino, 19-20 febbraio 2007). Frankfurt am Main: Peter Lang.
- CINQUE, Guglielmo (1999) *Adverbs and Functional Heads. A Cross-Linguistic Perspective*. Oxford: Oxford University Press.
- CORDIN, Patrizia (1993) "Dative Clitic Doubling in Trentino." In: A. Belletti (ed) *Syntactic Theory and the Dialects of Italy*. Torino: Rosenberg/Sellier, 130-154.
- CORDIN, Patrizia (1997) "Tense, mood and aspect in the verb." In: M. Maiden/M. Parry (eds) *The dialects of Italy*. London: Routledge, 87-98.
- CORDIN, Patrizia (2011) *Le costruzioni verbo-locativo in area romanza. Dallo spazio all'aspetto*. Berlin: de Gruyter.
- DEHÉ, Nicole/JACKENDOFF, Ray/MCINTYRE, Andrew/URBAN, Silke (eds) (2002) *Verb particle explorations*, Berlin/New York: de Gruyter.
- DUFRESNE, Monique/DUPUIS Fernande/TREMBLAY, Mireille (2003) "Preverbs and Particles in Old French." In: G. Booij/A. van Kemenade (eds), 33-60.
- DURANTE, Marcello (1981) *Dal latino all'italiano moderno. Saggio di storia linguistica e culturale*. Bologna: Zanichelli.
- FOLLI Raffaella/RAMCHAND Gillian (2005) "Prepositions and results in Italian and English: an Analysis from event decomposition." In: H. Verkuyl/H. de Swart/A. van Hout (eds) *Perspectives on aspect*. Dordrecht: Springer, 81-105.

- GRASSI, Corrado (2009) *Dizionario del dialetto di Montagne di Trento*. S. Michele all'Adige (Trento): Museo degli usi e costumi della gente trentina.
- GROFF, Lorenzo (1955) *Dizionario trentino-italiano*. Trento: Monauni.
- GSELL, Otto (1982) "Las rosas dattan ora – les röses da föra – le rose danno fuori: Verbalperiphrasen mit Ortsadverb im Rätoromanischen und im Italienischen." In: S. Heinz/U. Wandruszka (eds) *Fakten und Theorien. Festschrift für Helmut Stimm*. Tübingen: Narr, 71-85.
- HORROCKS, Geoffrey/STAVROU, Melita (2007) "Grammaticalized aspect and spatio-temporal culmination." *Lingua* 117, 605-644.
- IACOBINI, Claudio (2004) "Prefissazione." In: M. Grossmann/F. Rainer (edd), *La formazione delle parole in italiano*. Tübingen: Niemeyer, 98-163.
- IACOBINI, Claudio/MASINI, Francesca (2007) "The emergence of verb-particle constructions in Italian: locative and actional meanings." *Morphology* 16, 155-188.
- IACOBINI, Claudio/MASINI, Francesca (2009) "I verbi sintagmatici dell'italiano fra innovazione e persistenza: il ruolo dei dialetti." In: A. Cardinaletti/N. Munaro (eds), 115-135.
- ILIESCU, Maria/SILLER RUNGALDIER, Heidi/DANLER, Paul (eds) (2010) *Actes du XXV Congrès international de linguistique et philologie romanes*, Innsbruck, 3-8 septembre 2007, Voll. 7. Berlin: de Gruyter.
- JANSEN, Hans (2004) "La particella spaziale e il suo combinarsi con verbi di movimento nell'italiano contemporaneo." In: P. D'Achille (ed) *Generi, architetture e forme testuali*. Firenze: Cesati, 129-144.
- JEZEK, Elisabetta (2002) "Lo sfondamento di un confine tipologico. Il caso dei verbi complessi nell'italiano." In: P. Cordin/R. Franceschini/G. Held (eds) *Parallela 8. Lingue di confine, confini di fenomeni linguistici. Grenzsprachen. Grenzen von linguistischen Phänomenen*. Roma: Bulzoni, 289-306.
- JEZEK, Elisabetta/STRIK-LIEVERS, Francesca (2010) "Verbi sintagmatici in italiano antico e moderno. Un'analisi corpus-based." In: M. Iliescu/H. Siller Rungaldier/P. Danler (eds), 445-454.
- KING, Robert (2000) *The Lexical Bases of Grammatical Borrowing*. Amsterdam: Benjamins.
- KRAMER, Johannes (1987) "Tedeschismi e pseudo tedeschismi nel ladino e altrove." *Quaderni Patavini di Linguistica* 6, 9-30.
- MASINI, Francesca, (2006) "Diacronia dei verbi sintagmatici in italiano." *Archivio Glottologico Italiano* 91, 67-105.
- MATEU, Jaume/RIGAU, Gemma (2010) "Verb-particle constructions in Romance. A lexical-syntactic account." *Probus* 22, 241-269.
- PLANGG, Guntram (1980) "Le interferenze linguistiche: tedesco-ladino." *Mondo ladino* 3, 89-100.
- POLETTI, Cecilia (2000) *The higher functional field*. Oxford: Oxford University Press.
- POLETTI, Cecilia (2009) "I costrutti verbo + preposizione: l'interferenza tra veneto e italiano regionale." In: A. Cardinaletti/N. Munaro (eds), 155-172.

- ROHLFS, Gerhard (1969) *Grammatica storica della lingua italiana e dei suoi dialetti*. Vol. 3, *Sintassi e formazione delle parole*. Torino: Einaudi.
- SCHWARZE, Christoph (1985) "Uscire e andare fuori: struttura sintattica e semantica lessicale." In: A. Franchi De Bellis/L. Savoia (eds) *Sintassi e morfologia della lingua italiana d'uso. Teorie e applicazioni descrittive*. Atti del XVII Congresso internazionale di Studi della Società di Linguistica Italiana (Urbino 1983). Roma: Bulzoni, 355-371.
- SIMONE, Raffaele (1997) "Esistono verbi sintagmatici in italiano?" In: T. De Mauro/V. Lo Cascio (eds) *Lessico e grammatica*. Roma: Bulzoni, 155-170.
- SPIESS, Federico (1983) "L'unità lessicale composta di verbo e avverbio di luogo nei dialetti veneti." In: G. Holtus/M. Metzeltin (eds) *Linguistica e dialettologia veneta. Studi offerti a Manlio Cortelazzo dai colleghi stranieri*. Tübingen: Narr, 165-168.
- SPREAFICO, Lorenzo (2009) *Problemi di tipologia lessicale. I verbi di moto nello Standard Average European*. Roma: Bulzoni.
- TALMY, Leonard (1991) "Path to Realization: a Typology of Event Conflation." In: *Proceedings of the seventeenth annual meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistic Society, 480-519.
- TALMY, Leonard (2000) *Towards a cognitive semantics*, Cambridge, MA, MIT Press.
- TORTORA, Cristina (2002) "Romance Enclisis, Prepositions and Aspect." *Natural Language and Linguistic Theory* 20, 725-756.
- VICARIO, Federico (1997) *I verbi analitici in friulano*. Milano: FrancoAngeli.
- VIGOLO, Maria Teresa (2007) "Indicatori spaziali in dialetti alto-veneti e in cadornino." In: R. Maschi/N. Penello/P. Rizzolatti (eds) *Miscellanea di studi linguistici offerti a Laura Vanelli da amici e allievi padovani*. Udine: Forum, 353-363.
- VINCENT, Nigel (1999) "The evolution of c-structure: prepositions and PPs from Indo-European to Romance." *Linguistics* 37, 1111-1154.

#### Abstract

#### FROM VERBAL PREFIXES TO DIRECTION/RESULT MARKERS IN ROMANCE

In this paper I analyze verb-locative constructions in Romance. Even though not allowed in standard Romance languages, which have maintained and codified the classical Latin prefix system, these constructions are widely attested in non standard varieties, that are scarcely (or not at all) regularized. In this paper I deal in particular with a northern Italian variety, Trentino, where locatives, combining with some classes of verbs (unaccusative verbs and transitive activity verbs) can express not only concrete direction and metaphorical direction, but also aspect (the result of an activity or the progress of a process); in other words, they can express an abstract, more grammaticalized feature of direction. In fact, Trentino verb-locative constructions can express a specific phase of the development of an event, often the result (end-point) of an activity, or the start-point, or the progress, or the intensity of the activity itself.

I argue that in Trentino (and in other similar dialectal varieties) locatives can function as grammatical markers both for arguments (in combination with motion verbs) and for aspect (in combination with verbs involving the feature of an abstract path). In a more general context, I note that locatives in combination of verbs function as goal/result markers in those varieties that tend to spell out markers of functional elements, frequently generating a repetition (doubling) of the same feature.

Finally, I compare Romance and Germanic constructions, noting that diachronic, grammatical and typological differences (with particular reference to Talmy's distinction between verb-framed languages and satellite languages) confirm the independence of Romance verb-locative constructions.

#### Povzetek

### OD GLAGOLSKIH PREDPON DO OZNAČEVALCEV SMERI/REZULTATA V ROMANSKIH JEZIKIH

V pričujočem članku analiziramo analitične konstrukcije glagola in prostega prostorskega morfema v romanskih jezikih. Čeprav te konstrukcije niso sprejemljive v knjižnih romanskih jezikih, ki so ohranili in normirali predponski sistem klasične latinščine, ostajajo široko rabljene v večinoma skoraj povsem nenormiranih neknjižnih zvrsteh. V članku se ukvarjamo zlasti s severnoitalijanskim trentinskim govorom, kjer lahko prosti prostorski morfemi v kombinacijami z nekaterimi vrstami glagolov (netožilniški glagoli in prehodni glagoli dejanja) izražajo ne samo konkretno oziroma metaforično smer, ampak tudi glagolski aspekt (rezultat dejanja oziroma trajajoči proces). Povedano drugače, izrazijo lahko abstraktne, bolj gramatikalizirane pomene smeri. V trentinskem govoru lahko omenjene analitične konstrukcije izražajo specifično fazo v razvoju dogodka, pogosto rezultat (zaključek) dejanja, njegov začetek ali razvoj, lahko tudi intenzivnost samega dejanja.

Zagovarjamo trditev, da v trentinskem govoru (in sorodnih dialektalnih različicah) prosti prostorski morfemi delujejo kot slovnični označevalci tako za argumente (ko se vežejo z glagoli premikanja) kot za glagolski aspekt (ko se vežejo z glagoli, ki vključujejo pomen abstraktne poti). Na splošno opažamo, da delujejo prosti prostorski morfemi v kombinaciji z glagoli kot označevalci cilja/rezultata v tistih govorih, ki težijo k uporabi označevalcev kot funkcionalnih elementov, ki vodijo pogosto v (redundantno) ponovitev iste pomenske poteze.

V zadnjem delu primerjamo konstrukcije v romanskih in germanskih jezikih, kjer lahko na podlagi diahronih, slovničnih in tipoloških razlik (tu se posebej opiramo na Talmyjevo razlikovanje med jeziki, osredinjenimi na glagol, in jeziki, ki se opirajo na glagolske satelite) potrdimo, da gre v primeru glagolskih konstrukcij s prostim prostorskim morfemom za neodvisno nastale romanske strukture.



## MORPHOMES AS A LEVEL OF REPRESENTATION CAPTURE UNITY OF EXPONENCE ACROSS THE INFLECTION-DERIVATION DIVIDE\*\*

The premise of an inferential-realisation theory of inflectional morphology is that given an input consisting of a word's lexical identity (expressible for example by a lexical index,  $\lambda$ ) and a set of morphosyntactic properties,  $\sigma$ , provided by the syntax the inflectional morphology generates an output form of the word,  $\phi_{\text{WORD}(\lambda,\sigma)}$ , presumably expressed as an underlying phonological form. This output form  $\phi_{\text{WORD}(\lambda,\sigma)}$  is typically derived via some set of (morpho)phonological operations,  $\Phi_1, \Phi_2 \dots \Phi_n$ , performed on a lexical stem form,  $\phi_{\text{STEM}(\lambda,k)}$ , which is one of potentially many such stem forms belonging to lexeme  $\lambda$  (and where the choice of  $k$  will depend on  $\lambda$  and  $\sigma$ ). In sum, the inflectional morphology provides a mapping as in (1).

$$(1) \text{ The inflectional morphology: } \{\lambda, \sigma\} \rightarrow \phi_{\text{WORD}(\lambda,\sigma)} = \Phi_1 \cdot \Phi_2 \cdot \dots \cdot \Phi_n \cdot \phi_{\text{STEM}(\lambda,k)}$$

In (1) the composition of the operations  $\Phi_1, \Phi_2 \dots \Phi_n$  and stem form  $\phi_{\text{STEM}(\lambda,k)}$  is indicated as ' $\cdot$ '. Depending on the framework employed the nature of this composition ' $\cdot$ ' may vary. In a rule based framework for example ' $\cdot$ ' will typically translate into the assignment of  $\phi_{\text{STEM}(\lambda,k)}$  as the argument of the function composition of  $\Phi_1 \dots \Phi_n$ , so that (1) is implemented as (2).

(2) The inflectional morphology (rule-based framework):

$$\{\lambda, \sigma\} \rightarrow \phi_{\text{WORD}(\lambda,\sigma)} \Phi_1 (\Phi_2 (\dots (\Phi_n (\phi_{\text{STEM}(\lambda,k)})) \dots))$$

In other frameworks, such as a constraint based optimisation approach like Optimality Theory (Prince/Smolensky 1993/2004), the operation ' $\cdot$ ' need not be implemented as function composition, but it must be implemented in some way.<sup>1</sup>

---

\* *Author's address:* University of Queensland, School of Languages and Comparative Cultural Studies, Gordon Greenwood Building, St Lucia QLD 4072, Australia. Email: e.round@uq.edu.au

\*\* This research was supported in part by NSF Grant BCS 844550 and an Australian Research Council grant 'Isolation, Insularity and Change in Island Populations: An Interdisciplinary Study of Aboriginal Cultural Patterns in the Gulf of Carpentaria'.

<sup>1</sup> Presumably the function ' $\cdot$ ' will be non-commutative in any implementation, so that in the general case  $\Phi_i \cdot \Phi_j$  does not equal  $\Phi_j \cdot \Phi_i$ . In a rule-based implementation, this corresponds to the notion that rule ordering matters.

Turning to the nature of inflectional systems themselves, it is an empirical fact that inflectional systems of natural languages abound with cases in which two or more inflected forms of a lexeme are identical, that is, for two morphosyntactic property sets  $\sigma$  and  $\tau$ , where  $\sigma \neq \tau$ , it happens that  $\phi_{\text{WORD}(\lambda, \sigma)} = \phi_{\text{WORD}(\lambda, \tau)}$ . Multiple instances are apparent in the partial paradigm of *stelan*, a class IV strong verb of Old English shown (3).

(3) Partial paradigm of Old English *stelan* ‘to steal’

		1sg	2sg	3sg	1, 2, 3pl
Indicative	present	stele	stil(e)st	stil(e)þ	stelap
	past	stæl	stæle	stæl	stælon
Subjunctive	present	stele	stele	stele	stelen
	past	stæle	stæle	stæle	stælen

A central concern of inferential–realisational theories of inflection has been to provide a cogent account of such identities of exponence, and various tools have been developed to that end, more on which below. Notwithstanding achievements within inflection, Beard (1995) stresses the fact that identities of exponence also occur in derivation and more significantly that they can be found spanning the divide between derivation and inflection. An example of the latter can be seen in the wide distribution of the exponent *-ing* in Modern English (4).

(4) Some derivational and inflectional uses of English *-ing* (after Beard 1995:33)

Derivation:	resultative nominal	He brought his cuttings in.
	subjective adjective	It was a very cutting remark.
Inflection:	progressive aspect	The boy is cutting flowers.

In recent work on the Australian language Kayardild I have argued for a specific kind of analysis to account for a set of distinctive, complex patterns of identity of exponence within the inflectional system (Round 2009, in prep.). In this paper I examine the sharing of exponents in Kayardild across the derivation–inflection divide, and argue that the analysis developed earlier for Kayardild inflection generalises, so as to extend to derivation as well. The paper is organised as follows. Section 1 introduces the treatment of identities of exponence in an inferential–realisation approach to inflectional morphology, including an interpretation of the proposal in Round (2009) to account for certain problems raised by Kayardild. Section 2 examines Kayardild inflection and section 3 expands the discussion to Kayardild derivation. Conclusions are offered in section 4.

## 1. IDENTITY OF EXPONENCE AND ITS ANALYSIS IN INFLECTION<sup>2</sup>

Identities of exponence can occur at the level of entire words or, when words are morphologically complex, at the level of individual parts of words. For example in (3) above the past indicative 1st and 3rd person singular forms, *stæ*l, share a complete word form, while the subjunctive plural forms *stelen* (present) and *stælen* (past) share the inflectional suffix *-en*. Within inferential–realisational morphology there are two primary analytical tools for capturing identities such as these, and both can be applied to both whole-word and part-word identities. For ease of exposition in §1 I will assume a rule-based implementation of the inflectional analysis.<sup>2</sup>

The first technique is underspecification in the definition of the domain of a mapping, so that it relates more than one set of morphosyntactic properties  $\sigma$  to a single stem selection  $\Phi_{\text{STEM}(\lambda, k)}$  or stem modification  $\Phi_n$ . This technique will assign an identity of exponence to a NATURAL class of morphosyntactic properties, by virtue of the fact that morphosyntactic properties are typically arranged – on independent, semantic grounds – into values and features which pick out such natural classes. For example the statement (5a) makes no reference to person and number; it is underspecified for both, and hence picks out the natural class of all present subjunctive strong verbs. Similarly, the statement in (5b) is underspecified for person, tense and for the strong/weak verb contrast and hence picks out the natural class of all plural subjunctive verbs.

(5) Statements in the analysis of Old English inflection:

- a. Present subjunctive strong verbs take a certain stem (such as *stel-* in (3))
- b. Plural subjunctive verbs take the suffix *-en*

A second technique is the Rule of Referral (Zwicky 1985). This is more a powerful technique insofar as it can express identities of exponence across NON-NATURAL classes of morphosyntactic features. As originally formulated, Rules of Referral state for some set  $\sigma$  of morphosyntactic properties, that  $\Phi_{\text{WORD}(\lambda, \sigma)} =_{\text{def}} \Phi_{\text{WORD}(\lambda, \tau)}$ , where  $\tau$  ( $\neq \sigma$ ) is some other set of morphosyntactic properties. For example, if our analysis already states that the past indicative 1st person singular form of *stelan* is *stæ*l, then a Rule of Referral can state that the past indicative 3rd person singular is identical to the past indicative 1st person singular, thereby capturing the fact that the 3rd person singular form is also *stæ*l. Significantly, there is nothing in the nature of a Rule of Referral which requires  $\sigma$  and  $\tau$  to form a natural class.

Rules of Referral can also be relativised to specific parts of the derivation of a word's form (Stump 1993). Casting this notion in general terms, suppose that some part of the derivation is responsible for mapping from  $\sigma$  to just some part of  $\Phi_{\text{WORD}(\lambda, \sigma)}$  – it might be responsible only for selecting the stem  $\Phi_{\text{STEM}(\lambda, k)}$ , or for assigning one of the modifications  $\Phi_n$  for example. A Rule of Referral relating the

---

<sup>2</sup> This can be done without loss of generality, as there is nothing about the techniques which inherently ties them to a rule-based implementation.

realisation of  $\sigma$  to the realisation of  $\tau$  could then be constrained so as to have an effect only within this part of the derivation. As such, it will not in general force  $\phi_{\text{WORD}(\lambda,\sigma)}$  and  $\phi_{\text{WORD}(\lambda,\tau)}$  to be identical, but it will force some part of the realisations of  $\sigma$  to  $\tau$  to be identical.

The two main tools for capturing identities of exponence introduced so far are underspecification and Rules of Referral. A third alternative for capturing identity of exponence across non-natural classes of morphosyntactic properties was proposed by Aronoff (1994) under the rubric of the ‘morphome’. Aronoff (1994) in fact introduces two fundamentally different kinds of morphome. The first, which will be of little interest here is a feature belonging to a lexeme, one which influences how the lexeme and its associated sets of morphosyntactic properties are realised within the inflectional system. A classic example of this first kind of morphome is an inflectional class feature or declension feature. This kind of morphome serves to divide up the LEXICON into various parts whose inflectional pattern shares some significant similarity of form. The second kind of morphome which Aronoff introduces is more like a non-natural class of morphosyntactic property sets. These morphomes divide up the range of MORPHOSYNTACTIC PROPERTY SETS into groups whose realisation shares some significant similarity of form. In Latin for example the same stem form, termed the ‘third stem’ is always found across a group of morphosyntactic property sets which defies reduction into a natural class, appearing in the supine, the past participle and the future participle. If a given third stem appears in one of these forms it appears identically in all of them, no matter how irregular it might otherwise be, and should a lexeme be defective and lack one of these forms it will lack them all (see further Aronoff 1994:32–59). The Latin third stem is neither a specific form nor a natural morphosyntactic class, but rather is pattern of IDENTITY of exponence.

The interpretation which I would like to place on this second kind of morphome is that it functions in the inflectional morphology as an intermediate representation, located between the input  $[\lambda,\sigma]$  and the output  $\phi_{\text{WORD}(\lambda,\sigma)}$ , and moreover that it is a linguistically significant intermediate representation, and not merely a representation which arises as an artefact of a particular implementation of the analysis (such as the partially derived forms that arise as intermediate representations in a serial, rule-based derivations). The idea is that any two representations which are distinct from one another at the morphomic level will be realised as distinct in the output, and conversely that any two word forms  $\phi_{\text{WORD}(\lambda,\sigma)}$  and  $\phi_{\text{WORD}(\lambda,\tau)}$  which are (non-accidentally<sup>3</sup>) identical will have identical representations at the morphomic level. Similarly, identical PARTS of words’ realisations should be expressed by identity of

---

<sup>3</sup> Any theory of identity of exponence in inflection will encounter the issue of deciding between instances of identity that are linguistically significant and instances which are accidental. While this is an important issue, its resolution is orthogonal to the concerns in this paper, and the formalism I introduce here does not force any particular resolution on the analyst. For evidence of the psychological reality of morphomes as evidenced by historical change, see Maiden (2005).

PARTS of the morphomic representation. The morphomic level therefore expresses IDENTITIES of form, without expressing the forms themselves.

The model of inflectional morphology advocated is shown in (6).

(6) The inflectional morphology, with a morphomic level

$$\begin{aligned} [\lambda, \sigma] &\rightarrow \mu_{\text{WORD}(\lambda, \sigma)} = M_1 \cdot M_2 \cdot \dots \cdot M_m \cdot \mu_{\text{STEM}(\lambda, k)} \\ &\rightarrow \phi_{\text{WORD}(\lambda, \sigma)} = \Phi_1 \cdot \Phi_2 \cdot \dots \cdot \Phi_n \cdot \phi_{\text{STEM}(\lambda, k)} \end{aligned}$$

In (6) the input lexeme and morphosyntactic property set  $[\lambda, \sigma]$  maps to a morphomic representation  $\mu_{\text{WORD}(\lambda, \sigma)}$ , which in turn maps onto the underlying-phonological output,  $\phi_{\text{WORD}(\lambda, \sigma)}$ . The morphomic representation  $\mu_{\text{WORD}(\lambda, \sigma)}$  is composed of individual morphomic operations  $M_1, M_2 \dots M_m$  together with a lexical stem element  $\mu_{\text{STEM}(\lambda, k)}$  to be discussed further in §3. To be clear, there is no assumption that the morphomic operations  $M_1, M_2 \dots M_m$  map in a one-to-one fashion onto the phonological operations  $\Phi_1, \Phi_2 \dots \Phi_n$  (note the different subscripts for  $M_m$  and  $\Phi_n$ ). Just what this view of inflection corresponds to in empirical terms will become apparent when we consider the inflectional system of Kayardild.

## 2. KAYARDILD INFLECTION UNDER A MORPHOMIC ANALYSIS

Kayardild is a Tangkic (non-Pama Nyungan) language of northern Australia with a complex inflectional system characterised by affix stacking and pervasive identities in inflectional exponence. The language is described and analysed in a descriptive grammar by Evans (1995) and in a doctoral dissertation by Round (2009). According to Round (2009), Kayardild has two morphological word classes, nominal and verbal, and an inflectional system organised in terms of six morphosyntactic features which are privative and in general multi-valued, which is to say a word may be unspecified for a feature, or be specified for one of its several values, as listed in (7).

(7) a. case	23 values
b. number	2 values
c. thematic TAM	14 values
d. athematic TAM	11 values
e. negation	1 value
f. complementisation	2 values

The case system includes many semantically rich case values (Evans 1995), leading to a large number of values. The tense/aspect/mood (TAM) system is built on two features (Round 2009), one of which is termed ‘modal case’ in Evans (1995). Like the case system, the dual TAM system is semantically rich and has a large number of values for each feature.

For the purposes of the present discussion what is particularly interesting about Kayardild is the way in which identities of form are shared. To begin with, identities in exponence get shared across the paradigms of lexemes which have different parts

of speech: nominal case suffixes are often identical to verbal tense/aspect/mood (TAM<sup>4</sup>) suffixes. An example is shown in (8), where the oblique case and hortative TAM are both realised by the suffix /-ijca/. Kayardild words are shown in their underlying phonological forms in the middle columns and in (surface) orthographic form at the far right.

(8)	gloss	stem	oblique case	hortative TAM	
a.	'animal'	jaɭpuɭ	jaɭpuɭ-ijca	—	<i>yarbuthinja</i>
b.	'one'	waɭŋi:c	waɭŋi:c-ijca	—	<i>warnŋiijinja</i>
c.	'reed sp.'	kurkaŋ	kurkaŋ-ijca	—	<i>kurkanginja</i>
d.	'to leave'	ɭanaɭ	—	ɭanaɭ-ijca	<i>danathinja</i>
e.	'to go'	warac	—	warac-ijca	<i>warrajinja</i>
f.	'to go-NEG'	waranaŋ	—	waranaŋ-ijca	<i>warrananginja</i>

The derivations of the word forms in (8a–f) are shown respectively in (9a–f), with an explanation to follow.

(9)	input		morphomic level → output level
a.	{YARBUTH,σ}	→	M <sub>T</sub> ·M <sub>OBL</sub> ·μ <sub>STEM</sub> (YARBUTH,1)
		→	Φ <sub>T</sub> ·Φ <sub>-inja</sub> /jaɭpuɭ/
b.	{WARNGIIL,σ}	→	M <sub>T</sub> ·M <sub>OBL</sub> ·μ <sub>STEM</sub> (WARNGIIL,1)
		→	Φ <sub>T</sub> ·Φ <sub>-inja</sub> /waɭŋi:c/
c.	{KURKANG,σ}	→	M <sub>T</sub> ·M <sub>OBL</sub> ·μ <sub>STEM</sub> (KURKANG,1)
		→	Φ <sub>T</sub> ·Φ <sub>-inja</sub> /kuɭkaŋ/
d.	{DANATH,τ}	→	M <sub>T</sub> ·M <sub>OBL</sub> ·μ <sub>STEM</sub> (DANATH,1)
		→	Φ <sub>T</sub> ·Φ <sub>-inja</sub> /ɭanaɭ/
e.	{WARRAJ,τ}	→	M <sub>T</sub> ·M <sub>OBL</sub> ·μ <sub>STEM</sub> (WARRAJ,1)
		→	Φ <sub>T</sub> ·Φ <sub>-inja</sub> /warac/
f.	{WARRAJ,v}	→	M <sub>T</sub> ·M <sub>OBL</sub> ·M <sub>NEG</sub> ·μ <sub>STEM</sub> (WARRAJ,1)
		→	Φ <sub>T</sub> ·Φ <sub>-inja</sub> ·Φ <sub>-rang</sub> /warac/

where:

σ = {case:oblique}

τ = {TAM:hortative}

v = {TAM:hortative, +negative}

<sup>4</sup> The TAM properties referred to in this section and the next are thematic TAM values. Similar issues arise with respect to the athematic TAM properties and the complementisation feature.

At the morphomic level all of the representations in (9a–f) share two morphomic operations  $M_T$  and  $M_{OBL}$ <sup>5</sup> and by virtue of that they all share some part of their output form.  $M_{OBL}$  is the realisation at the morphomic level of both {case:oblique} and [TAM:hortative].  $M_T$  is an obligatory morphological operation that applies to every word in Kayardild (see Round 2009:150–65). At the phonological level,  $M_{OBL}$  is realised as the operation  $\Phi_{-inja}$  which adds the suffix /-inja/ to a stem.<sup>6</sup>  $M_T$  is realised as  $\Phi_T$ , an operation which suffixes /-tá/, /-ka/ or /-a/ to a stem, or which effects no change, depending on the phonology of the stem. In the case of the stems in (8), all of which end in /a/,  $\Phi_T$  effects no change and hence the final piece of phonology in all of the forms is the suffix /-inja/.

It is significant in (9) that the sharing of exponence between oblique case forms and hortative TAM forms does not rely on it being possible for a given lexeme to inflect for both. That is, lexical nominal stems in Kayardild cannot inflect for hortative TAM and nor can lexical verbal stems inflect for oblique case. As such, the identities in exponence between (8a–c) on the one hand and (8d–f) on the other could not be captured by Rules of Referral such as (10a, b).

(10) Impossible Rules of Referral in Kayardild:

- a.       \* REALISATION-OF[ $\lambda$ ,{case:oblique}]  
          =def REALISATION-OF[ $\lambda$ ,{ TAM:hortative }]
- b.       \* REALISATION-OF [ $\lambda$ ,{ TAM:hortative }]  
          =def REALISATION-OF[ $\lambda$ ,{case:oblique}]

This is no accident. A limitation of Rules of Referral is that they cannot be used to express identities in the forms that realise two sets of morphosyntactic properties  $\sigma$  and  $\tau$  if  $\sigma$  and  $\tau$  are not compatible with the same lexemes. The empirical evidence from Kayardild indicates that natural language morphologies do not in general face the same restriction. The morphomic approach advocated here has the desirable ability to capture the identities of exponence that are present in cases like (8).

The situation illustrated in (8) is by no means an isolated case. According to Round (2009) the inflectional system of Kayardild overtly realises fifty-three mor-

<sup>5</sup> Following a convention in Round (2009) based in turn on Evans (1995) a morphomic operation is named after any morphosyntactic case feature which it realises.  $M_{OBL}$  is labelled after case:oblique, which it realises in (8/9a–c). The morphomic operation  $M_{NEG}$  in (9f) never realises a case feature, and so is labelled after another morphosyntactic feature which it does realise, in this instance negative:yes.  $M_T$  is a ‘termination’ which is a semantically empty piece of morphology.

<sup>6</sup> In Kayardild there are two major kinds of phonological juncture which can separate morphs, indicated as ‘<sup>-</sup>’ and ‘<sup>+</sup>’ in underlying phonological forms, and in general the surface realisation of /a-b/ and /a+b/ will not be the same. The determinants of which junctures appear where in Kayardild are complex but systematic. For reasons of space they will not be considered here, but see Round (2009, forthc.) for discussion. Essentially, an account of the distribution of the two junctures requires an embellishment of morphomic representations which in the end furnishes an additional source of support for the general morphomic approach advocated here.

phosyntactic feature values.<sup>7</sup> Of those, twenty eight have a morphomic realisation consisting of a single morphomic operation,  $M_n$ , which they share with at least one other feature-value. Some of these cases are illustrated in (11) and (12). (Note that in (11) and (12)  $M_T$  has a non-null realisation in the output, because  $\Phi_T$  adds the suffix /-a/ to a stem ending in /i/.)

(11)	gloss	stem	locative case	immediate TAM	
a.	‘animal’	jaɭpuɬ	jaɭpuɬ+ki-a	—	<i>yarbuthiya</i>
b.	‘one’	waɳɳi:c	waɳɳi:c+ki-a	—	<i>warnɡiiɳiya</i>
c.	‘reed sp.’	kurkaŋ	kurkaŋ+ki-a	—	<i>kurkangiya</i>
d.	‘to leave’	ʃanaɬ	—	ʃanat+ki-a	<i>danathiya</i>
e.	‘to go’	warac	—	warac+ki-a	<i>warraɳiya</i>
f.	‘to go-NEG’	waranaŋ	—	waranaŋ+ki-a	<i>warranangiya</i>
(12)	gloss	stem	privative case	actual TAM & +neg	
a.	‘animal’	jaɭpuɬ	jaɭpuɬ-wari-a	—	<i>yarbuyarriya</i>
b.	‘one’	waɳɳi:c	waɳɳi:c-wari-a	—	<i>warnɡiiyarriya</i>
c.	‘reed sp.’	kurkaŋ	kurkaŋ-wari-a	—	<i>kurkawarriya</i>
d.	‘to leave’	ʃanaɬ	—	ʃanat+wari-a	<i>danatharriya</i>
e.	‘to go’	warac	—	warac+wari-a	<i>warraɳarriya</i>

Another relevant phenomenon in Kayardild is the realisation of certain morphosyntactic properties by two morphomic operations,  $M_1$  and  $M_2$ , both of which serve individually to realise other morphosyntactic properties. Examples are shown in (13) with derivations in (14). In (13/14) the operation  $\Phi_T$  adds the suffix /-a/ to stems ending in /n/ and /-a/ to stems ending in /i/.

(13)	stem, gloss	properties	output	
a.	ʃanaɬ	{TAM:continuous}	ʃanaɬ-n-ʃa	<i>dananda</i>
b.	‘to leave’	{TAM:actual, +neg}	ʃanaɬ+wari-a	<i>danatharriya</i>
c.		{TAM:nonveridical}	ʃanaɬ-n-wari-a	<i>dananmarriya</i>
d.	warac	{TAM:continuous}	warac-n-ʃa	<i>warranda</i>
e.	‘to go’	{TAM:actual, +neg}	warac+wari-a	<i>warraɳarriya</i>
f.		{TAM:nonveridical}	warac-n-wari-a	<i>warranmarriya</i>
(14)	input	morphomic level → output level		
a.	{WARRAJ,σ}	→	$M_T \cdot M_N \cdot \mu_{STEM-WARRAJ,1}$	
		→	$\Phi_T \cdot \Phi_{-n} / \text{warac} /$	
b.	{WARRAJ,τ}	→	$M_T \cdot M_{PRIV} \cdot \mu_{STEM-WARRAJ,1}$	
		→	$\Phi_T \cdot \Phi_{-warri} / \text{warac} /$	
c.	{WARRAJ,v}	→	$M_T \cdot M_{PRIV} \cdot M_N \cdot \mu_{STEM-WARRAJ,1}$	
		→	$\Phi_T \cdot \Phi_{-warri} \cdot \Phi_{-n} / \text{warac} /$	

<sup>7</sup> These are: 23 case values, 2 number values, 11 ‘athematic’ TAM values, 14 ‘thematic’ TAM values, 2 complementisation values, and neg:yes.



- d. [DANATH,σ] →  $M_T \cdot M_N \cdot \mu_{STEM \cdot DANATH,1}$   
→  $\Phi_T \cdot \Phi_{-n} / \{ \text{ana}\} /$
- e. [DANATH,τ] →  $MT \cdot M_{PRIV} \cdot \mu_{STEM \cdot DANATH,1}$   
→  $\Phi_T \cdot \Phi_{-warri} / \{ \text{ana}\} /$
- f. [DANATH,v] →  $M_T \cdot M_{PRIV} \cdot M_N \cdot \mu_{STEM \cdot DANATH,1}$   
→  $\Phi_T \cdot \Phi_{-warri} \cdot \Phi_{-n} / \{ \text{ana}\} /$

where:

σ = {TAM:continuous}

τ = {TAM:actual, +negative}

v = {TAM:nonveridical}

The point of interest here is that the forms (13c,f) show non-accidental identities of form with both (13a,d) and (13b,e). These identities are captured at the morphomic level in terms of (14c,f) sharing  $M_N$  with (14a,d) and  $M_{PRIV}$  with (14b,e). As was the case earlier, these identities of exponence are not expressible with Rules of Referral. The failure of Rules of Referral in this instance occurs because lexemes inflected in (14) with the morphosyntactic property set  $v$  simultaneously share identities of exponence with the same lexemes inflected with  $\sigma$  AND with the same lexemes inflected with  $\tau$ . This is true even though no one lexeme in Kayardild will ever be associated simultaneously with both  $\sigma$  and  $\tau$ . That is to say, a Rule of Referral such as (15) will fail because it attempts to refer on its right hand side to an ill-defined set of morphosyntactic properties ( $\sigma \cup \tau$ ).

(15) Impossible Rule of Referral in Kayardild:

\* REALISATION-OF[ $\lambda$ , {TAM:nonveridical}]

=def REALISATION-OF[ $\lambda$ ,

{TAM:continuous}  $\cup$  {TAM:actual, +neg}]

The pattern illustrated in (13) is also well-attested in Kayardild. Of the fifty-one overtly realised morphosyntactic feature values in Kayardild, ten are realised by multiple morphomic operations of which at least one figures in the realisation of a different feature value. An example which is parallel to (13) is shown in (16). In (16)  $\Phi_T$  adds the suffix /- $\{a\}$ / to stems ending in / $n$ / and leaves to stems ending in / $a$ / unaffected.

(16) stem, gloss	properties	output	
a. $\{ \text{ana}\}$	{TAM:continuous}	$\{ \text{ana}\} \cdot n \cdot \{ \text{a}\}$	dananda
b. 'to leave'	{TAM:precondition}	$\{ \text{ana}\} + \eta \text{arpa}$	danatharrba
c.	{TAM:antecedent}	$\{ \text{ana}\} \cdot n \cdot \eta \text{arpa}$	danann Garrba
d. warac	{TAM:continuous}	warac $\cdot n \cdot \{ \text{a}\}$	warranda
e. 'to go'	{TAM:precondition}	warac $+ \eta \text{arpa}$	warrajarrba
f.	{TAM:antecedent}	warac $\cdot n \cdot \eta \text{arpa}$	warrann Garrba

### 3. DERIVATION AND MORPHOMIC REPRESENTATIONS OF LEXICAL STEMS

We can now examine the nature of the morphomic element  $\mu_{\text{STEM}(\lambda,k)}$ , which is the morphomic representation of a lexical stem. I propose that  $\mu_{\text{STEM}(\lambda,k)}$  has an internal structure as in (17). That is,  $\mu_{\text{STEM}(\lambda,k)}$  itself is composed of zero or more morphomic operations plus the phonological form of some root  $\phi_{\text{ROOT}}$ .

- (17) The internal structure of  $\mu_{\text{STEM}(\lambda,k)}$ , the morphomic representation of a lexical stem  

$$\mu_{\text{STEM}(\lambda,k)} = (M_1 \cdot M_2 \cdot \dots \cdot M_n) \phi_{\text{ROOT}} \quad ; \text{ where } (\dots) \text{ indicates optionality}$$

By extending the representational device of morphomic operations,  $M_n$ , into the stem it will be possible to capture identities of form which span the divide between derivation and inflection. The need to capture such identities is argued for prominently by Beard (1995) and remarked upon by Stump (2001:203–7), while Aronoff (1994:34,126–7) makes explicit reference to morphemes which unite derivational and inflectional categories. Likewise, the extension of essentially the same morphological architecture from inflection to derivation (irrespective of whether this results in shared exponents) has been advocated by Bauer (1997) with respect to paradigmatic arrangements of suffixes, Booij (1997) with respect to stem selection, and Stump (2001:252–60) with respect to fundamental principles of paradigm structure. The representation in (17) thus provides a means of achieving ends which have been identified as appealing in prior research. Let us turn then to the empirical evidence in Kayardild in support of (17).

In the inflectional system of Kayardild the realisation of its fifty-one overtly-marked morphosyntactic feature values requires a inventory of just twenty-eight morphomic operations,  $M_n$ . Of those twenty-eight, fifteen figure in the realisation of more than one feature value. Furthermore fifteen – though not exactly the same fifteen – are employed derivationally.<sup>8</sup> Examples of these identities of form across the derivation–inflection divide will follow below. In setting out the examples I will not be concerned with how a morphologically complex stem is derived<sup>9</sup> in the lexicon but rather with how its form is represented, with a particular focus on how that representation captures aspects of shared exponence with inflected words.

The morphomic operation  $M_{\text{ASSOC}}$  is used in Kayardild as the realisation of the morphosyntactic feature value case:associative. It is in turn realised by the phonological operation  $\Phi_{-rnurru}$  which adds the suffix /-uru/ to its stem. Example (18a)

<sup>8</sup> The distinction in Kayardild between inflectional and derivational morphology is clear cut: an inflectional marker on a nominal stem will appear on all other words in a noun phrase that that nominal occupies, and an inflectional marker on a verbal stem will appear on all other verbal words in the same clause (other than those at a deeper level of clausal embedding); neither of these conditions holds of derivational markers (Evans 1995:88–89).

<sup>9</sup> For argument's sake, we can suppose that a derivationally complex stem form is a realisation of a root index  $\rho$  plus some set of derivational properties  $\delta$ , with a mapping from  $\{\rho, \delta\}$  in much the way that inflection maps from  $\{\lambda, \sigma\}$ . The forms to which  $\{\rho, \delta\}$  map are the concern of this section.

shows the inflectional derivation the lexeme WUMBURUNG ‘spear thrower’ inflected with associative case.

(18a) Input	Morphomic level → output level
a. i. [WUMBURUNG, {case:assoc}]	→ $M_T \cdot M_{ASSOC} \cdot \mu_{STEM(WUMBURUNG,1)}$
ii.	= $M_T \cdot M_{ASSOC} / wumpu\text{ɬ}uŋ/$
iii.	→ $\Phi_T \cdot \Phi_{-muru} / wumpu\text{ɬ}uŋ/$
iv.	= $wumpu\text{ɬ}uŋ\text{-}ŋuru\text{-}a$ <i>wumpurnurruwa</i>

In (18a.i) the lexical index WUMBURUNG and morphosyntactic properties {case:associative} map to a morphomic representation, including the morphomic representation of the lexical stem,  $\mu_{STEM(WUMBURUNG,1)}$ . The operation  $M_{ASSOC}$  is the realisation of case:associative and  $M_T$  is the usual operation which appears on all Kayardild words. Because the lexical stem of WUMBURUNG is morphologically simple,  $\mu_{STEM(WUMBURUNG,1)}$  expands in (a.ii) to a simple root /wumpuɬuŋ/. The representation in (a.ii) then in turn maps to a phonological representation in (a.iii) which expands to (a.iv).

Example (b) shows a derivationally complex lexeme KURNDURNURRU ‘woman with young child’ which is based on the root /kuŋɬuŋ/, where the simple lexeme KURNDUNG means ‘chest’.

(18b)	Input	Morphomic level → output level
i.	[KURNDURNURRU, {∅}]	→ $M_T \cdot \mu_{STEM(KURNDURNURRU,1)}$
ii.		= $M_T \cdot M_{ASSOC} / kuŋ\text{ɬ}uŋ/$
iii.		→ $\Phi_T \cdot \Phi_{-muru} / kuŋ\text{ɬ}uŋ/$
iv.		= $kuŋ\text{ɬ}uŋ\text{-}ŋuru\text{-}a$ <i>kurndurnurruwa</i>

Line (18b.i) maps from the lexeme KURNDURNURRU with its empty set of associated morphosyntactic properties<sup>10</sup> to a morphomic representation. In (18b.i) the only morphomic operation shown is the usual  $M_T$ . In (18b.ii) the lexical representation  $\mu_{STEM(kurndurnurru,1)}$  is expanded out. My proposal is that KURNDURNURRU is represented morphomically as  $M_{ASSOC} / kuŋɬuŋ/$ . This allows the framework to capture the non-accidental identity of form between (18a.ii) above where  $M_{ASSOC}$  realised an inflectional feature and (b.ii) where  $M_{ASSOC}$  (and its phonological realisation /-ŋuru/) is part of a morphologically complex lexical stem. Because (18a.ii) and (18b.ii) have parallel representations at the morphomic level (where IDENTITY, but not content, of form is captured) the mappings and expansions in (18a.iii–iv) and (18b.iii–iv) are also entirely parallel.

<sup>10</sup> Here as in Round (2009) I treat a lack of case marking as following from the absence of a case value in a word’s morphosyntactic properties.

In (19a) the morphomic operation  $M_{DU}$  is the realisation of number:dual, which is realised on the simple lexeme DUN ‘husband’. In (19b) the complex lexeme MUNKIYARRNG ‘whale’ is realised with an empty set of morphosyntactic properties. The simple lexeme MUN means ‘buttocks’ (whence MUNKIYARRNG alludes to a whale’s bifurcated tail). In (19b.ii) the lexical representation  $\mu_{STEM}(MUNKIYARRNG,1)$  expands to  $M_{DU}/mun/$  so that the morphomic representations in (19a.ii) and (19b.ii) are parallel, thereby capturing the identities of form and causing the mappings and expansions in (19a.iii–iv) and (19b.iii–iv) to be commensurate also.

(19)	Input	Morphomic level → output level
a. i.	{DUN,[number:dual]}	→ $M_T \cdot M_{DU} \cdot \mu_{STEM}(DUN,1)$
ii.		= $M_T \cdot M_{DU} / \{ \text{tun} /$
iii.		→ $\Phi_T \cdot \Phi_{kiyarrng} / \{ \text{tun} /$
iv.		= $\{ \text{tun} + \text{kiarrŋ} - \text{ka} \quad \text{dunkiyarrngka}$
b. i.	{MUNKIYARRNG, {∅}}	→ $M_T \cdot \mu_{STEM}(MUNKIYARRNG,1)$
ii.		= $M_T \cdot M_{DU} / \{ \text{mun} /$
iii.		→ $\Phi_T \cdot \Phi_{kiyarrng} / \{ \text{mun} /$
iv.		= $\text{mun} + \text{kiarrŋ} - \text{ka} \quad \text{munkiyarrngka}$

Example (20) is inflectionally more complex than (18) and (19). In (20a) the nominal lexeme WUMBURUNG ‘spear thrower’ inflects for case:donative, and also for TAM:actual. The TAM feature comes into play because the donative is one of Kayardild’s case values, termed ‘verbal(ising) case’ (Evans 1995, 2003) or ‘thematic case’ (Round 2009), which lead to a nominal inflecting also for TAM/negation. Thus in (20a.i) we see case:donative mapping to  $M_{TH} \cdot M_{DON}$ ,<sup>11</sup> TAM:actual mapping to  $M_{ACT}$ , and of course  $M_T$ . In (a.iii) the morphomic operations  $M_T \cdot M_{ACT}$  map cumulatively to  $\Phi_{-a}$ , which adds the suffix /-a/ to its stem;  $M_{TH}$  maps to  $\Phi_j$  and  $M_{DON}$  to  $\Phi_{-wu}$ .

(20a)	Input	Morphomic level → output level
i.	{WUMBURUNG, {case:donative, TAM:actual}}	→ $M_T \cdot M_{ACT} \cdot M_{TH} \cdot M_{DON} \cdot \mu_{STEM}(WUMBURUNG,1)$
ii.		= $M_T \cdot M_{ACT} \cdot M_{TH} \cdot M_{DON} / \{ \text{wumpu} \text{ɟuŋ} /$
iii.		→ $\Phi_{-a} \cdot \Phi_j \cdot \Phi_{-wu} / \{ \text{wumpu} \text{ɟuŋ} /$
iv.		= $\text{wumpu} \text{ɟuŋ} - \text{wu} - \text{c} - \text{a} \quad \text{wumburuwuja}$

The morphologically complex verb KAWUJ is built on the lexical root KANG ‘speech’ and means ‘to speak to OBJ, where OBJ is a kinsman to whom one is permitted by law to speak’. In (20b) KAWUJ is inflected with{TAM:actual} that is, with a proper subset of the inflectional features involved in (20a). In (20b.i) the input maps to

<sup>11</sup> The subscript th on  $M_{TH}$  stands for ‘thematic’, hence the name ‘thematic case’ in Round (2009).

$M_{ACT}$  which realises the inflectional features,  $M_T$  as per usual, and to the lexical representation  $\mu_{STEM(KAWUJ,1)}$  which then expands in (20b.ii) into  $M_{TH} \cdot M_{DON} / ka\eta /$ , which I propose as the morphomic representation of KAWUJ. Consequently the morphomic representations in (20a.ii) and (20b.ii) are parallel, capturing the identities of form involved and leading the mappings and expansions in (20a.iii–iv) and (20b.iii–iv) to be commensurate also.

(20b)	Input	Morphomic level → output level
i.	[KAWUJ, {TAM:actual}]	
		→ $M_T \cdot M_{ACT} \cdot \mu_{STEM(KAWUJ,1)}$
ii.		= $M_T \cdot M_{ACT} \cdot M_{TH} \cdot M_{DON} / ka\eta /$
iii.		→ $\Phi_{-a} \cdot \Phi_{-j} \cdot \Phi_{-wu} / ka\eta /$
iv.		= ka $\eta$ -wu-c-a <i>kawuja</i>

#### 4. CONCLUSION

In Round (2009, in prep.) I propose the existence of an intermediate, ‘morphomic’ level of representation, which plays a significant role in the inferential–realisational analysis of inflection. Specifically it makes possible the formal expression of certain kinds of identities of exponence, common in Kayardild, which involve non-natural classes of morphosynactic properties, but which are not readily expressible by Rules of Referral. In this paper I have shown that rather similar identities of form appear across the derivation–inflection divide in Kayardild, and that the same devices employed in the analysis of Kayardild inflection can be generalised profitably to derivation. Doing so requires recognising that lexical stems have not only a phonological form but a morphomic form too. Once that move is made, the morphomic level of representation assumes the generalised role of expressing patterns of identities of form, without expressing the forms themselves, across inflection and derivation.

#### References

- ARONOFF, Mark (1994) *Morphology by itself: stems and inflectional classes*. Cambridge, MA: MIT Press.
- BAUER, Laurie (1997) “Derivational paradigms.” In: G. Booij/J. van Marle (eds) *Yearbook of Morphology 1996*, Dordrecht: Kluwer, 243–56.
- BEARD, Robert (1995) *Lexeme-morpheme base morphology*. Albany, NY: SUNY Press.
- BOOIJ, Geert (1997) “Autonomous morphology and paradigmatic relations.” In: G. Booij/J. van Marle (eds) *Yearbook of Morphology 1996*, Dordrecht: Kluwer, 35–53.
- EVANS, Nicholas (1995) *A grammar of Kayardild: with historical-comparative notes on Tangkic*. Berlin: Mouton de Gruyter.
- EVANS, Nicholas (2003) “Typologies of agreement: some problems from Kayardild.” *Transactions of the Philological Society* 101:203–34.

- MAIDEN, Martin (2005) "Morphological autonomy and diachrony." *Yearbook of Morphology 2004*, New York: Springer, 137–75.
- ROUND, Erich (2009) *Kayardild Morphology, Phonology and Morphosyntax*. Ph.D. dissertation, Yale University, New Haven.
- ROUND, Erich (in prep.) *Kayardild Inflection*. Oxford: Oxford University Press.
- STUMP, Gregory (1993) "On Rules of Referral." *Language* 69: 449–79.
- STUMP, Gregory (2001) *Inflectional morphology: a theory of paradigm structure*. Cambridge: Cambridge University Press.
- ZWICKZY, Arnold (1985) "How to describe inflection." In: M. Niepokuj, M. Van Clay, V. Nikiforidou/D. Feder (eds) *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, BLS, 372–86.

#### Abstract

#### MORPHOMES AS A LEVEL OF REPRESENTATION CAPTURE UNITY OF EXPONENCE ACROSS THE INFLECTION-DERIVATION DIVIDE

Inferential–realisational analyses formalise a language's inflectional morphology in terms of a mapping on the one side from a lexical index and set of morphosyntactic properties to on the other side a phonological form. Round (2009) has argued that the Australian language Kayardild requires the postulation of an intermediate level of representation, identified with Aronoff's (1994) notion of a 'morphome'. This morphomic level serves to express patterns of identities of exponence abstracted away from the actual forms of exponents and its use makes possible the expression of certain identities of form which defy expression by means of Rules of Referral (Zwicky 1985, Stump 1993). This paper considers identities of form that span the inflection-derivation divide in Kayardild and shows that they too are coherently captured by assuming that a morphomic level of representation is present. A consequence is that lexical stems must possess a morphomic representation in addition to their representations on other levels.

#### Povzetek

#### MORFOMI KOT PREDSTAVITVENA RAVNINA ZAJAMEJO ENOTNOST IZRAZNIH PODOB NE GLEDE NA RAZKOL MED OBLIKOSLOVJEM IN BESEDOTVORJEM

Inferenčno-uresničitvena analiza oblikoslovje določenega jezika formalizira v obliki preslikav iz seznama leksikalnih enot in množice morfosintaktičnih lastnosti na eni strani v fonološko obliko na drugi strani. V Round (2009) smo zagovarjali trditev, da avstralski jezik kayardild zahteva vpeljavo vmesne predstavitvene ravnine, ki ustreza Aronoffovemu (1994) pojmu »morfoma«. S pomočjo te morfomske ravnine lahko izražamo vzorce izrazne enakosti (ang. *identity of exponence*) neodvisno od dejanskih izraznih oblik (ang. *forms of exponents*), prav tako ta raven omogoča izražanje določenih enakosti oblik, ki jih ni mogoče zajeti s *pravili podajanja* (Zwicky 1985, Stump 1993). Pričujoči članek obravnava enakosti oblik, ki v jeziku kayardild segajo preko ločnice med oblikoslovjem in besedotvorjem, in dokazuje, da je tudi te enakosti moč sistematično zajeti, pod pogojem, da je prisotna morfomska predstavitvena ravnina. Posledica je, da morajo imeti leksikalne podstave tudi morfomsko predstavitev in ne samo predstavitev na drugih ravninah.

## MOROCCAN ARABIC BORROWED CIRCUMFIX FROM BERBER: INVESTIGATING MORPHOLOGICAL CATEGORIES IN A LANGUAGE CONTACT SITUATION

### 1. INTRODUCTION

Moroccan Arabic (MA, hereafter) has a bound derivational noun circumfix /ta- . . -t/. This circumfix productively attaches to base nouns to derive various types of abstract nouns. This /ta- . . -t/ circumfix is unique because it occurs in no other dialect of Arabic (Abdel-Massih 2009, Bergman 2005, Erwin 1963). There is little question that this morpheme has been borrowed from Berber into Moroccan Arabic, as has been noted by some of the earliest analyses of the language (c.f. Guay 1918, Colon 1945).

This circumfix is highly productive on native MA noun stems but not productive on borrowed Berber stems (which are rare in MA). This pattern of productivity is taken to be evidence in support of direct borrowing of morphology (c.f. Steinkruger and Seifart 2009) and against a theory where borrowed morphology enters a language as part of unanalyzed complex forms which later spread to native stems (c.f. Thomason and Kaufman 1988; Thomason 2001). Furthermore, it challenges the principle of a “borrowability hierarchy” (c.f. Haugen 1950) where lexical morphemes are borrowed before grammatical morphemes. Additionally, the prefixal portion of the MA circumfix, *ta-*, is a complex (presumably unanalyzed) form from the Berber /t-/ feminine + /a-/ absolute state. Moreover, the morpheme in MA has been borrowed as a derivational morpheme while the primary functions of the donor morphemes in Berber are inflectional.

This case of a borrowed circumfix is examined as an instance which does not conform to traditional typologies of language change through contact patterns and how the boundaries between morphological categories (i.e. inflectional/derivational) are transformed in a language contact situation.

In section 2, the form and functions of the circumfix in MA and Berber are detailed. Section 3 discusses the MA borrowed circumfix in relation to current theories about the typology of borrowed morphology. Section 4 discusses the findings of this paper and their importance with respect to generalizations about the typology of language change through contact and the boundaries of morphology.

---

\* *Author's address:* University of Colorado at Boulder, Department of Linguistics, Hellems 290, 295 UCB, Boulder CO 80309-0295, USA. Email: georgia.weissman@colorado.edu

## 2. BORROWED CIRCUMFIX IN MOROCCAN ARABIC FROM BERBER

The form and function of the circumfixes in MA and Berber are outlined in this section.

### 2.1. Form and Function of the MA Circumfix

A circumfix can be defined as a discontinuous morpheme that surrounds a base, with one part of the morpheme acting as a prefix and the other part of the morpheme acting as a suffix (Harris and Xu 2006). The circumfix in MA takes the form /ta- . -t/ when attaching to base nouns. This marker is considered a circumfix because neither the prefixal part /ta- nor the suffixal part /-t/ occur with a base alone to derive the same function.

The diffused MA circumfix has the function of deriving abstract nouns from base nouns. There are three main types of abstract nouns derived with the circumfix: nouns of profession, nouns of human characteristics, and general abstract nouns.

The first type of abstract noun that this morpheme derives in MA is a noun of occupation with the sense of “the profession/art/occupation of x” from a base noun of occupation with the meaning of “x.” The forms in (1) show the base lexical items which refer to an occupation and then the derived nouns with the circumfix which refer to ‘the art/profession’ of that occupation.

#### (1) Moroccan Arabic

Base Noun	Gloss	Derived Abstract Noun	Gloss
bənnay	‘mason’	tabənnayət	‘profession/art of masonry’
buwwab	‘doorman’	tabuwwabət	‘profession/art of doorman’
kotbi	‘bookseller’	takotbiyət	‘profession/art of bookseller’
bramli	‘barrel-maker’	tabramlit	‘profession/art of barrelmaking’

(Harrell et al. 1963)

The second type of abstract noun the circumfix can derive in MA is a noun of quality/characteristics of people from base nouns with related meanings. Examples of these base and derived nouns are illustrated in (2).

#### (2) Moroccan Arabic

Base Noun	Gloss	Derived Abstract Noun	Gloss
kəddab	‘liar’	takəddabət	‘habitual lying’
ʔadami	‘polite’	taʔadamit	‘politeness’

(Harrell 1962, Harrell et al. 1963)



Finally, general abstract nouns can be derived from concrete base nouns, as illustrated in (3), below:

(3) Moroccan Arabic

Base Noun	Gloss	Derived Abstract Noun	Gloss
dərri	‘child’	tadərrit	‘childhood’
šgaŋ	‘ill-fated (person)’	tašguŋiyət	‘bad luck’

Note that the underlying form of the circumfix is considered here to be /ta- . . -t/. This form surfaces when attached to stems that end in vowels, as shown above in the words *taḡadami* ‘politeness’ < *ḡadami* ‘polite’ and *tabramlit* ‘profession of barrel-maker’ < *bramli* ‘barrel-maker’. However, when the stem ends in a consonant an epenthetic schwa [ ] is inserted between the stem and suffixal portion of the circumfix. This can be seen above in the words *takəddabət* ‘habitual liar’ < *kəddab* ‘liar’ and *tabənnayət* ‘profession of masonry’ < *bənnay* ‘mason’. Thus, the circumfix has conditioned allomorphy depending on the form of the stem it attaches to.

Importantly, all of the base nouns given in Tables (1)-(3) are MA native stems. Thus, the circumfix is highly productive on native stems.

## 2.2. Form and Function of the Berber Circumfix

The Berber languages have a circumfix /t- . . -t/ which has several functions. First, the Berber circumfix is used in the formation of grammatically feminine nouns in the singular. Second, the circumfix is used to form abstract nouns. Furthermore, the circumfix can form a noun of single unit from a collective noun. Finally, the circumfix can form diminutive nouns from masculine nouns.

The following examples illustrate the functions of the /t- . . -t/ circumfix the Moroccan Berber languages. The /a/ morpheme is a nominal marker of absolute state, as is explained in section 3.1.1, below. The most predominant function of the circumfix in the Berber languages is to mark nouns as grammatically feminine, as shown below in Tashelhit:

(4) Tashelhit (Southern Morocco)

Feminine Nouns – Absolute State	Gloss
t-a-šlh-t	‘Tashelhit, female speaker’
t-a-frux-t	‘girl’
t-a-fuk-t	‘sun’

(Dell and Elmedlaoui 2002: 7, 17, 26)

The nouns in (4), above are shown with the feminine circumfix in the absolute state, marked by /a-/. However, there is another way to inflect these nouns, which is for annex state. The same nouns are given below in (5) in the annex state, which is marked by a zero morpheme with the feminine circumfix:

(5) Tashelhit (Southern Morocco)

Feminine Nouns – Annex State	Gloss
t-šlh-t	‘Tashelhit, female speaker’
t-frux-t	‘girl’
t-fuk-t	‘sun’

(Ibid)

Note that the /a-/ is a marker of noun state in Berber. The feminine circumfix does occur in Berber without initial /a-/ when it is inflected for annex state.

The function of marking feminine gender applies to animate nouns, as in (5)a. and b. where the noun stem without the circumfix corresponds to the masculine meaning (c.f. *a-šlh* ‘Tashelhit, male speaker’; *a-frux* ‘boy’). Additionally, this function applies to inanimate nouns where the circumfix conveys grammatical gender, as in (5)c.

The second function of the /t- . . -t/ circumfix in Berber is to derive abstract nouns. Often, these nouns are derived from verb stems. The data below is from Tamazight:

(6) Tamazight (Central Morocco)

Abstract Nouns	Gloss
t-alum-t	‘blame’
t-ažallit-t	‘praying’
t-ikki-t	‘passing’

(Abdel-Massih 1971: 96)

The examples below in (7) illustrate the third function of the circumfix, which is to derive count nouns from collective nouns. The data below illustrates this process in Tamazight:

(7) Tamazight (Central Morocco)

Collective Noun	Gloss	Derived Noun of Unity	Gloss
ažalim	‘onions’	t-ažalim-t	‘an onion’
zznbuŕ	‘oranges’	t-azznbuŕ-t	‘an orange’
xizzu	‘carrots’	t-axizzu-t	‘a carrot’

(Abdel-Massih 1971: 118)

Finally, the circumfix is used as a derivational morpheme to create diminutive (or augmentative) nouns.

(8) Tashelhit (Southern Morocco)

Base Noun	Gloss	Diminutive	Gloss
a. udm	'face'	t-udm-t	'little face'
b. afus	'hand'	t-afus-t	'baby hand'

(Dell and Elmedlaoui 2002: 26, Kossmann 2008: 76)

In (8), a grammatically masculine noun *udm* 'face' can take the feminine circumfix to derive a diminutive *t-udm-t* 'little face'. This function also works essentially in reverse whereby grammatically feminine nouns with the circumfix can be derived into an augmentative. This is illustrated below in (9). A grammatically feminine noun *t-aryal-t* 'basket', which takes the circumfix, can be derived into an augmentative by removing the circumfix to create *aryal* 'large basket'.

(9) Tashelhit (Southern Morocco)

Augmentative	Gloss	Feminine Noun	Gloss
aryal	'large basket'	t-aryal-t	'basket'
amar	'large beard'	t-amar-t	'beard'

(Dell and Elmedlaoui 2002: 26, Kossmann in press: 76)

### 2.3. Comparing the Circumfix in MA and Berber

The two circumfixes differ in both form and function. First, the MA circumfix has been shown to be /ta- . -t/, while the circumfix in Berber is simply /t- . -t/. The difference is not great, but it reveals that this morpheme was borrowed as an unanalyzed complex form consisting of component inflectional morphemes, as discussed in the next section.

Another difference between the MA circumfix and the Berber circumfix is that while the latter occurs on both native and borrowed stems productively, the MA circumfix is mainly productive with MA native stems. This is odd considering it is a borrowed morpheme. This distributional difference is also examined in the next section. Finally, the functions of the MA and Berber circumfixes differ. While the circumfix (and its component parts) are inflectional morphemes in Berber, they have been analyzed and are used as a single derivational morpheme in MA. These facts explored in this paper make the MA circumfix an insightful instance of borrowed morphology and how the phonological, lexical, and semantic boundaries of morphology become transformed in situations of language contact.

Furthermore, it is worth noting that while the circumfix in Berber is a feminine or diminutive suffix, the morpheme becomes an abstract noun marker in Moroccan Arabic. In other words, there appears to have been a generalizing of the function of the morpheme from a feminine and abstract noun marker in Berber to just the abstract noun function in the receiving language. This is typical of morphology diffused from one language to another through contact.

### 3. PROBLEMS WITH THEORIES OF BORROWED MORPHOLOGY

There are many commonly held generalizations about what can be easily borrowed in a situation of language contact, as opposed to what is more rarely borrowed. For example, Haugen (1950) outlined a hierarchy of patterns based on the relative freedom of which they are borrowed. This typology generalized that “nouns are the most easily borrowed”, followed by other lexical categories, then affixes, then inflectional morphology, then sounds (224). He supported this hierarchy with evidence from loanword data.

More recently, studies of language contact reiterate this generalization about the hierarchy of borrowability. Thomason and Kaufman (1988) and Thomason (2001) outline predictions and typological generalizations as to which types of elements are borrowed in language contact contexts. They make the generalization that vocabulary is borrowed before structure (Thomason 2001: 69). This generalization is not supported by the borrowed circumfix in MA. This section outlines some of the features of the circumfix that do not conform to the much discussed typologies of borrowed morphology.

#### 3.1. Pattern of productivity

Note that the borrowed circumfix (in both the recipient and donor languages) is a bound morpheme – it always occurs attached to a base noun stem. As was shown in table (1)-(3), the borrowed circumfix is highly productive on native MA noun stems. However, the MA circumfix is not highly productive on borrowed Berber stems (since they are rare in MA). While the exact percentage of borrowed lexical items in MA from Berber has not been calculated to our knowledge, many researchers support the claim that “Berber borrows from MA more than the latter borrows from the former” due to both the prestige of MA since its speakers associate it with Standard Arabic and since Berber-MA bilinguals are usually L1 Berber speakers, not vice versa (Sadiqi 1997: 14, Ennaji 1997).

Statistics on the percentage of MA loanwords in Berber can be found. For example, the Northern Moroccan Berber language Tarafit has 51.7% of its lexicon consisting of loanwords, 41.7% of those loanwords originating from dialectal Arabic (Kossmann 2009: 197-8).

This pattern of productivity is taken to be evidence in support of direct borrowing of morphology (c.f. Steinkruger and Seifart 2009) and against a theory where borrowed morphology enters a language as part of unanalyzed complex forms which later spread to native stems (c.f. Thomason and Kaufman 1988; Thomason 2001). In addition, Thomason also claims that derivational affixes are borrowed through loanwords and later spread to native stems (2001: 70).

Evidence that the MA circumfix is an instance of direct borrowing of morphology comes from MA loanwords in Berber: a large number of MA loanwords in Berber are transformed to match the inflectional patterns of Berber; specifically, they are inflected for noun state and gender. The borrowed stems from MA that occur in Berber with Berber inflectional morphology can be found in MA with this Berber

inflectional morphology reanalyzed as a single derivational morpheme. There is no evidence that these native MA stems are being re-borrowed back into MA, which might be suggested if it is claimed that these forms were initially borrowed whole with the inflectional morphology unanalyzed, since these stems are not being semantically or phonologically transformed in any way and they occur prevalently in MA without the circumfix. Rather, it appears that the circumfix entered MA on these native stems and has been productively adopted by MA speakers and generalized as a derivational morpheme. This scenario of the MA circumfix as an instance of direct borrowing of morphology is explored in this section.

### 3.1.1. Noun State in Berber

Borrowed morphology does not always keep the same phonological shape when diffused from one language to another. However, finding evidence for why the shape of the morpheme has changed from one language to the other can give insight into what speakers do when encountered with imported morphological boundaries. There are two points of evidence which explain why the shape of the prefixal portion of the MA circumfix has been borrowed from Berber as /ta-/, rather than simply /t-/. First, the one of the two morphological forms of nouns in Berber, the absolute state, most frequently has initial vowel /a/ in the noun stem. Secondly, nouns borrowed into Berber from MA often take initial /a/.

All nouns in Berber occur in one of two morphological ‘states’: absolute (*état libre*) or annex (*état d’annexion*). A morphological change signals the two alternative forms of nouns, whose distribution is determined by syntactic position and other syntactic and morphological properties of the phrase.

The data given in (10) is a nominal paradigm for the noun stem a-drar ‘mountain’ in the Absolute and Annex states in Tashelhit Berber:

(10) Tashelhit (Southern Morocco)

*Nominal Paradigm: a-drar ‘mountain’*

	Masc. Sing	Fem. Sing	Masc. Pl	Fem. Pl.
Absolute	a-drar	t-a-drar-t	y-drar-n	t-i-drar-in
Annex	w-drar <sup>1</sup>	t-drar-t	w-drar-n	t-drar-in

(Dell and Elmedlaoui 2002: 29)

<sup>1</sup> Glides alternate with their vocalic counterparts when phonologically conditioned (Dell and Elmedlaoui 2002: 29-30).

Absolute state of masculine and feminine singular nouns is marked by initial /a-/ of the stem (Achab 2003: 2). In the annex state the stem is prefixed with either a /w-/ in the masculine or a zero morpheme in the feminine. Sometimes, stems in Berber have lexically specified initial vowels, such as /a/, /i/ or /u/, in which case absolute marker /a-/ is either assimilated to the initial vowel or deleted through hiatus resolution: adjacent or long vowels are not permitted in Northern Berber (Dell and Elmedlaoui 2002).

The addition of a schwa between the stem and the final /-t/ of the circumfix can be explained as epenthesis of a vowel, as evidenced by the fact that when a vowel-final stem takes the circumfix, no schwa is inserted (e.g. *bramli* ‘barrel-maker’, *ta-bramli-t* ‘profession of barrel-maker’).

The data on Berber noun classes absolute state and annex state begs the question: if the circumfix occurs in Berber around stems without any additional vowel, as in the feminine singular annex state form of the noun in (25) above, why was it not borrowed into MA in this form? The answer to this question lies in the distribution of noun state in Berber. The most important property of noun state in Berber, for our purposes, involves dependency implications. Specifically, nouns in absolute state can be independent – they can stand alone – while nouns in annex state are dependent in that they must occur and be interpreted with another element preceding it in a clause (Chaker 1988). This also follows Basset (1932): “l’état d’annexion [s’emploie] quand le nom, sujet du verbe, est placé après lui, quand il est complètement déterminatif, quand il est précédé d’une préposition: en somme quand la voyelle préradicale ou préformative tombe, elle tombe la ou le mot fait étroitement corps avec celui que le précède.» (173).

The examples below illustrate this point. A noun with the absolute marker /a-/, as given in (11), can be stated by itself.

- (11) Kabyle (Algeria)  
 t-a-zemmur-t  
 F-ABS-olive.tree-F  
 ‘An olive tree’

(Amina Mettouchi, p.c.)

However, with annex state morphology must occur post-verbally or following a preposition. For example, in (12)a., the noun *uzemmur* ‘an olive’ has an initial /w-/ marking annex state and it occurs after a quantifier *a eqqa* ‘a grain of’. Crucially, the noun *uzemmer* is ungrammatical if stated by itself, as in (12)b.

- (12) Kabyle (Algeria)  
 a. a eqqa u-zemmur  
 a.grain.of ANN-olive  
 ‘an olive’  
 b. \*u-zemmur  
 \*ANN-olive

(Amina Mettouchi, p.c.)

The circumfix in Berber is just /t...t/ to mark the feminine and most often occurs around stems with initial /a/ vowels inserted to mark absolute state. This would be the borrowed form because nouns in the absolute state can be stated independently. If this circumfix was borrowed with annex state morphology (without the vowel /a/), it would probably have to be borrowed with another syntactic element since annex state nouns in Berber always occur in a tight relationship with a preceding syntactic element. Since the stem in MA usually begins with a consonant, there has been a reanalysis of the Berber circumfix to have an initial /ta/ to mirror the shape of words with this circumfix in Berber.

### 3.1.2. /ta-/ as an unanalyzed complex form

The MA circumfix has the form /ta- . .-t/. However, the form of the Berber circumfix is simply /t- . .-t/ with no state morphemes. Thus, there is a slight difference in shape of the prefixal portion between the two morphemes, specifically an additional /a/ vowel which consistently appears in all the MA instances of the morpheme. However, a full vowel such as /a/ has never been described as being epenthetic in MA. The fact that the circumfix exists in MA with the /a/ has significance in regards to the nominal morphology of the donor language, Berber. Specifically, the /a/ vowel is a marker of absolute state in Berber, in nouns both with and without the feminine circumfix. The initial portion of the prefix in MA has been borrowed as an unanalyzed complex form /ta-/ with no meaningful component morphemes. It is not uncommon for languages to borrow a complex morphemes as an unanalyzed single morpheme, but since that this morpheme has been borrowed into MA from Berber without the conduit of Berber loanwords makes this an interesting fact. The circumfix in MA has been reanalyzed in the process of being borrowed without the conduit of large Berber loanwords in the language.

### 3.1.3. MA loanwords in Berber

Many MA lexical items have been borrowed into Berber. These loanwords often take Berber inflectional morphology. For example, in Tashelhit Berber, MA words are “berberized” (term from Aspinion 1953) by adding the absolute state morpheme /a-/ to the beginning of the stem, as shown below in (13).

(13) Tashelhit (Southern Morocco)

Borrowed Form	Gloss	Moroccan Arabic Form
abənnay	‘bricklayer’	< bənnay
axəbbaz	‘bread-maker’	< xəbbaz
aħəžžam	‘barber’	< həžžam

(Aspinion 1953: 9-36)

The loanwords from MA illustrated in (29) are borrowed into Tashelhit Berber with initial vowel /a-/ which marks absolute noun state. Loanwords can be fully

inflected for both noun state and gender, for example the feminine gender is marked with the /t...-t/ circumfix, as illustrated in (14):

(14) Tamazight (Central Morocco)

Masc. Form	Gloss	Fem Form	Gloss	Moroccan Arabic
a-xddam	‘worker M’	t-a-xddam-t	‘worker F’	xəddam
a-rddam	‘demolisher M’	t-a-rddam-t	‘demolisher F’	rəddam

(Abdel-Massih 1971: 92-5)

The forms in (14) show words borrowed from MA into Tamazight that not only have the inserted /a/ vowel in the masculine form, but take the feminine circumfix and the /a/ augment in the feminine form.

All the Arabic loanwords found in Berber given in (13) and illustrated in (15) can take the /ta- . . -t/ circumfix in MA as an abstract noun and occur in their un borrowed form (base noun).

(15) Moroccan Arabic

Base Noun	Gloss	Derived Abstract Noun	Gloss
bənnay	‘bricklayer’	tabənnayət	‘occupation of bricklayer’
xəbbaz	‘bread-maker’	taxəbbazət	‘occupation of bread-maker’
həžžam	‘barber’	tahəžžamət	‘occupation of barber’

(Harrell et al. 1963)

The pattern of borrowing evidenced here is one where lexical items are being borrowed from MA into Berber. In Berber, these nouns are inflected for noun state and gender like native Berber words. These native MA words can be found in MA with this Berber inflectional morphology reanalyzed as a single circumfix. Since the native MA base nouns are not transformed in any way it can be argued that the /ta- . . -t/ circumfix is being directly borrowed into MA from Berber using the native MA stems first as a conduit to later being used productively in the language.

**3.1.4. Direct Borrowing of Morphology**

The pattern that can be inferred for the MA circumfix is one where MA lexical items are diffused into Berber, the most common language contact scenario, then these stems are inflected to match the inflectional patterns of Berber, including noun state /a-/ and feminine gender /t...-t/. Finally, these nouns with Berber inflectional morphology are incorporated back into MA with the inflectional morphology components reanalyzed as a single derivational morpheme. Again, note that the circumfix in question is a bound morpheme in both donor and recipient languages.



Since the original nouns are native MA stems, and they have not been transformed in shape or semantics, the innovative morpheme component can be argued to be an example of directly borrowed bound morphology.

### 3.2. Borrowability Hierarchy

Thomason and Kaufman (1988) and Thomason (2001) outline predictions and typological generalizations as to which types of elements can be expected to be borrowed in various language contact situations. Their basic prediction is that vocabulary is borrowed before structure (Thomason 2001: 69). Another determinant for borrowing in language contact: “less tightly structured features are easier to borrow than features that fit into tightly integrated closed structures” (*ibid*: 69).

These statements are the most recent incarnation of what generalizations have been made earlier (c.f. Haugen 1950). However, more recent studies have shed light on the fact that these generalizations cannot be considered cross-linguistic. For example, Steinkruger and Seifart (2009) have cited numerous instances where bound derivational and inflectional morphology is borrowed without the borrowing of loanword stems (i.e. Bora morphology in Resígaro (North West Amazon) and Bisayan morphology in Chabacano de Zamboanga (Mindanao, Philippines)). Along with the evidence examined in this paper where the MA circumfix can be demonstrated to be an instance of where a bound morpheme is directly borrowed from Berber, generalization such as those made by Thomason et al. need to be reexamined. A theory of language contact where needs explain and describe direct borrowing of morphology in specific situations.

## 4. DISCUSSION

Typological studies of morphology in language contact have made generalization about how borrowed morphology gets diffused. The borrowed MA circumfix examined here does not follow these generalizations in numerous aspects. This is an unusual diffusion pattern for several reasons. Sociolinguistically, MA can be thought of as the prestige language, in relation to the Berber dialects spoken in the same region. This fact can be further supported, linguistically, by evidence that the majority of loanword borrowing goes in the direction of MA to Berber, rather than vice versa. However, this circumfix is an example of direct borrowing of morphology. Perhaps the typical notions of the order of what is borrowed in a language contact situation (i.e. borrowability hierarchy) and what happens in situations of language prestige in a language contact situation need to be reexamined.

Examining instances such as the MA circumfix, which do not conform to wide held generalizations about the formal and functional borders of borrowed morphology, is important.

## References

- ABDEL-MASSIH, Ernest T. (1971) *A Reference Grammar of Tamazight (Middle Atlas Berber)*. Ann Arbor: University of Michigan Press.
- ABDEL-MASSIH, Ernest T. (2009) *A Reference Grammar of Egyptian Arabic*. Georgetown: Georgetown University Press.
- ASPINION, Robert. (1953) *Apprenons le berbère*. Rabat: F. Moncho.
- BASSET, André. (1932) “Sur l'état d'annexion en berbère.” *Bulletin de la Société de Linguistique de Paris* XXXIII, 173-174.
- BERGMAN, Elizabeth. (2005) *Spoken Algerian Arabic*. Springfield, VA: Dunwoody.
- CHAKER, Salem. (1988) “Annexion (L'état d').” In: S. Chaker (ed.) *Encyclopedie berbère V*. Aix-en-Provence: Édisud, 689-695.
- COLON, Georges (1947) “Quelques ‘emprunts’ de morphèmes étrangers dans les parlers arabes occidentaux.” *Groupe linguistique d'Etudes Chamito-Sémitiques* IV.
- DELL, Francois and Mohamed ELMEDLAOUI. (2002) *Syllables in Tashlhiyt Berber and Moroccan Arabic*. The Netherlands: Kluwer.
- ENNAJI, Moha. (1997) “The Sociology of Berber: Change and Continuity.” *International Journal of the Sociology of Language* 123:1, 23-40.
- ERWIN, Wallace. (1963) *A Short Reference Grammar of Iraqi Arabic*. Georgetown: Georgetown University Press.
- GUAY, F. (1918) “La Forme féminine berbère à Sale.” *Archives Berbères* 3, 31-51.
- HARRELL, Richard. (1962) *A Short Reference Grammar of Moroccan Arabic*. Georgetown: Georgetown University Press.
- HARRELL, Richard. et al. (1963) *A Dictionary of Moroccan Arabic*. Georgetown: Georgetown University Press.
- HARRIS, Alice and Zheng XU. (2006) “Diachronic Morphological Typology.” In: *Encyclopedia of Language and Linguistics, Second Edition*, Ed. Keith Brown. Oxford: Elsevier, 509-515.
- HAUGEN, Einar (1950) “The Analysis of Linguistic Borrowing.” *Language* 26:2, 210-231.
- HEATH, Jeffrey. (2002) *Jewish and Muslim Dialects of Moroccan Arabic*. London: Routledge Curzon.
- KOSSMANN, Maarten G. (2008) “The collective in Berber and language contact.” In: M. Lafkioui/V. Brugnatelli (eds.) *Berber in Contact. Linguistic and Sociolinguistic Perspectives*. Köln: Rüdiger Köppe Verlag, 53-62.
- KOSSMANN, Maarten G. (2009) “Loanwords in Tarifiyt, a Berber language of Morocco.” In: Haspelmath, M/Tadmor, U (Eds.), *Loanwords in the World's Languages. A Comparative Handbook*. Berlin: De Gruyter Mouton, 191-214.
- KOSSMANN, Maarten G. In press. “Berber.” In: Zygmunt Frajzyngier/Erin Shay (eds.), in press: *The Afroasiatic Languages*. Cambridge: Cambridge University Press.
- SADIQI, Fatima (1997) “The Place of Berber in Morocco.” *International Journal of the Sociology of Language* 123:1, 7-22.
- STEINKRUGER, Patrick/Frank SEIFART (2009) “Transfer of Derivational Morphology without Borrowing of Stems: Resigaro (Arawakan, Peru) and Chabacano (Creole,

Phillipines).” Presented at *Morphologies in Contact*. University of Bremen, Germany, October 2009.

THOMASON, Sarah/Terence KAUFMAN. (1988) *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.

THOMASON, Sarah (2001) *Language Contact: An Introduction*. Washington D.C.: Georgetown University Press.

#### Abstract

### MOROCCAN ARABIC BORROWED CIRCUMFIX FROM BERBER: INVESTIGATING MORPHOLOGICAL CATEGORIES IN A LANGUAGE CONTACT SITUATION

Moroccan Arabic (MA) has a derivational noun circumfix /ta- . .-t/ that is borrowed from the neighboring Berber languages. This circumfix is highly productive on native MA noun stems but not productive on borrowed Berber stems (which are rare in MA). This pattern of productivity is taken to be evidence in support of direct borrowing of morphology (c.f. Steinkruger and Seifart 2009) and against a theory where borrowed morphology enters a language as part of unanalyzed complex forms which later spread to native stems (c.f. Thomason and Kaufman 1988; Thomason 2001); furthermore, it challenges the principle of a “borrowability hierarchy” (c.f. Haugen 1950) where lexical morphemes are borrowed before grammatical morphemes. Additionally, the prefixal portion of the MA circumfix, *ta-*, is a complex (presumably unanalyzed) form from the Berber /t-/ feminine + /a-/ absolute state. Moreover, the morpheme in MA has been borrowed as a derivational morpheme while the primary functions of the donor morphemes in Berber are inflectional.

This case of a borrowed circumfix is examined as an instance which does not conform to traditional typologies of language change through contact patterns and how the boundaries between morphological categories (i.e. inflectional/derivational) are transformed in a language contact situation.

## Povzetek

### BERBERSKI CIRKUMFIKS V MAROŠKI ARABŠČINI: RAZISKOVANJE MORFOLOŠKIH KATEGORIJ V NA PRIMERU JEZIKOV V STIKU

Maroška arabščina (MA) pozna besedotvorni cirkumfiks /ta-...-t/, ki je izposojen iz sosednjih berberskih govorov. Ta cirkumfiks je zelo produktiven na maroškoarabskih samostalniških podstavah, a neproduktiven na izposojenih berberskih podstavah, ki so sicer v MA redke. Ta produkcijski vzorec je služil kot dokaz za teorijo neposrednega izposojanja morfologije (cf. Steinkruger in Seifart 2009) in za zavrnitev teorije, po kateri izposojena morfologija v jezik vstopa kot del nemotiviranih tvorjenih oblik, ki lahko sicer kasneje predstavljajo podstave za domače besedotvorne procese (cf. Thomason in Kaufman 1988; Thomason 2001). Ta vzorec nadalje postavlja pod vprašaj načelo »izposojevalne hierarhije« (cf. Haugen 1950), v skladu s katerim se leksikalni morfemi izposojajo pred gramatikalnimi. Poleg tega je predponski del cirkumfiksa v MA, *ta-*, tvorjena (najverjetneje nemotivirana) oblika iz berberskih morfemov /t-/ za ženski spol in /a-/ za absolutno stanje. Še več, morfem je bil v MA izposojen kot besedotvorni morfem, medtem ko je glavna funkcija obeh izvirnih morfemov v berberščini oblikoslovna.

Omenjeni izposojeni cirkumfiks v članku analiziramo kot primer, ki se ne ujema s tradicionalnimi tipologijami jezikovnih sprememb preko različnih vrst stikov, in primer, kako se lahko zaradi stikov med jeziki premaknejo meje med morfološkimi kategorijami (v našem primeru med oblikoslovjem in besedotvorjem).

## CREATIVE COMPETENCE IN BORROWINGS: WORDS OF JAPANESE ORIGIN IN SLOVENE

### 1. INTRODUCTION

‘Karaoke’<sup>1</sup>, ‘anime’<sup>2</sup>, ‘shiitake’<sup>3</sup>. These are words of Japanese origin borrowed in many European languages today. In the case of English (or even German, French etc.) these words basically retain their original appearance, as they are employed in the context in the recipient language, since the recipient language has no noun inflection. In case of Slovene, it is different. Here are some published examples of these words in context<sup>4</sup>:

(1) *Je že res, da živimo v dobi karaok, ... [Dnevnik 2005-03-04]*

“It’s true that we live in the karaoke era, ...”

(2) *Kot sem rekel: mange bere staro in mlado. Isto velja za gledanje animejev.*

[*Joker* 2006-08-01]

“As I said: mangas are read by old and young. The same holds true for watching animes.”

(3) *5 šitakov [Svet in ljudje 2004-02]*

“5 shiitakes” (in a recipe)

Though the original forms of these loanwords all end in the same vowel –*e* and all are used in the Slovene sentences in the genitive case, ‘karaoke’ is used in sentence (1) in the feminine plural, ‘anime’ in sentence (2) in the masculine plural, and ‘shi-

---

\* *Author’s address*: Filozofska fakulteta, Oddelek za azijske in afriške študije, Aškerčeva 2, 1000 Ljubljana, Slovenia. Email: chikako.bucar@guest.arnes.si

<sup>1</sup> [mass noun] a form of entertainment, offered typically by bars and clubs, in which people take turns to sing popular songs into a microphone over pre-recorded backing tracks. - ORIGIN 1970s: from Japanese, literally ‘empty orchestra’. (Oxford Dictionary of English 2003)

<sup>2</sup> [mass noun] Japanese film and television animation, typically having a science-fiction theme and sometimes including violent or explicitly sexual material. - ORIGIN 1980s: Japanese. (Oxford Dictionary of English 2003)

<sup>3</sup> (also shitake, shiitake mushroom) an edible mushroom which grows on fallen timber, cultivated in Japan and China. *Lentinus edodes*, family Pleurotaceae, class Hymenomycetes. - ORIGIN late 19<sup>th</sup> cent.: from Japanese, from *shii*, denoting a kind of oak, + *take* ‘mushroom’. (Oxford Dictionary of English 2003)

<sup>4</sup> Example sentences are taken from *Fidaplus*, the Slovene referential corpus. In the square bracket are the source (periodicals) and the date of appearance as indicated in the corpus.

itake (*šitake*)’ in sentence (3) in the masculine singular. If we try to trace these words back to the form of nominative and genitive singular in Slovene which is usually used for dictionary entries, they would be: *karaoka -e*; *anime -ja*; and *šitak -i*.<sup>5</sup>

There are several reasons and different motivations for the assignment of number and gender to loanwords before users of the recipient language arrive at agreement for certain linguistic rules. In the following sections, the grammatical characteristics of the source language (Japanese) and the recipient language (Slovene) in relation to the process of borrowing are reviewed (sections 2 and 3), and possible rules of gender and number assignment to loanwords of Japanese origin are given (section 4). The degree of integration into the recipient language for the loanwords is further verified through derivation of new words (section 5). Based on the findings in the case of borrowings of Japanese origin into Slovene, the creative competence of speakers is discussed.

## JAPANESE

### 2.1 Noun morphology

Japanese has no category of number or of gender. Nouns may be used to denote singular or plural in one and the same form according to context. In addition, Japanese nouns (as well as adjectives) have no case inflection, but the syntactic relationships to verbs are expressed by the use of postpositional case particles. These characteristics may be observed in the following example sentences:

(4) *Kinoo*                      *ryoori*                      *no*                      *hon*                      *wo*                      *katta.*  
 yesterday                      cooking                      GEN                      book                      ACC                      buy-PAST  
 “I bought a cookbook yesterday.”

(5) *Kinoo*                      *tosyokan no*                      *hon*                      *wo*                      *seiri sita.*  
 yesterday                      cooking GEN                      book                      ACC                      put in order-PAST  
 “I put the books in the library in order yesterday.”

The noun *hon*, ‘book’, appears in both sentences in the same form, but it is translated as “one book” in example (4) and plural “books” in example (5). Of course, sentence (4) may also mean “I bought cookbooks yesterday.”, if the speaker loves cooking and often buys cookbooks. In the case of example (5), the noun *hon* is difficult to understand as a singular “book”, since there are always many books in any library.

The voice relationship of the noun to the verb is expressed by the postpositional case particle, in the above cases *wo* to denote the accusative (or the direct object) case. The same noun may be accompanied by some other case particle, as in the following examples, and the form of the noun remains unchanged.

---

<sup>5</sup> The dictionary entries for the Slovene language is usually given in the order of nominative singular – genitive singular, e.g. *lipa -e* ‘linden’ (feminine), *kot -a* ‘corner’ (masculine), *mleko -a* ‘milk’ (neuter).

- (6) *Ryoori no hon ga hosii.*  
 cooking GEN book NOM desired  
 “I want a cookbook (I want cookbooks).”
- (7) *Ryoori no hon ni hana wo soete okutta.*  
 cooking GEN book DAT flower ACC accompany-CONT send-PAST  
 “I added flowers to the cookbook and sent them.”  
 “I sent a cookbook together with flowers.”

## 2.2 Phonology

Phonologically speaking, Japanese is a syllabic language with basically open syllables which means that all syllables end in a vowel, with only two exceptions. Two kinds of moraic consonants /Q/<sup>6</sup> and /N/<sup>7</sup> may close a syllable, but within the Japanese phonological system, both are counted as mora-building. A mora is the standard rhythmic unit in Japanese.<sup>8</sup> Theoretically, the structures of all possible morae in Japanese are: V, CV, Q, and N. Since the standard Japanese language has five vowels, a mora (and a syllable) of the type V or CV ends in one of the five vowels: *a*, *e*, *i*, *o* or *u*. Regular Japanese nouns (and all other words) consist of one, two, three or more morae. For example, most simple nouns take the following form: *te* ‘hand’, *ha* ‘tooth’, *asi* ‘foot’, *kutu* ‘shoe’, *atama* ‘head’, *otoko* ‘man’.

All five vowels in Japanese may be either short or long. Shibatani (1990) explains:

Long vowels contrast with short vowels, e.g. *su* ‘vinegar’: *suu* ‘number’, *to* ‘door’ : *too* ‘ten’. A long vowel cannot simply be considered a single segment, for it is counted as two moras, and, furthermore, because pitch drop occurs in the middle of a long vowel, e.g. *sūu* (HL) ‘number’, just as in any two mora sequence, e.g. *mīso* (HL) ‘bean paste’. (Shibatani 1990: 161-162)

Another important phonological characteristic of Japanese is the accentual system, the so-called pitch accent. In modern standard Japanese, each mora is spoken

<sup>6</sup> The phoneme /Q/ is a non-nasal consonant followed by a homorganic consonant of the following syllable, realized phonetically as a prolongation of consonantal articulation e.g. *yappari* /jaQpari/ [jappari] ‘also’, *assari* /aQsari/ [assari] ‘plainly’. (Shibatani 1990: 158, 167-8) Both of the example words consist of 4-morae, since the mora-building /Q/ counts as one mora.

<sup>7</sup> The phoneme /N/ is a nasal that closes a syllable, and unlike the non-nasal /Q/, it can end a word. In careful pronunciation it is close to the uvular nasal [N]. In less careful speech it is realized as a nasalized version of the preceding vowel, e.g. *hon* /hoN/ [hō̃] ‘book’. (Shibatani 1990: 169) Further examples: *tonbo* /toNbo/ [tombo] ‘dragonfly’; *ringo* /riNgo/ [ringo] ‘apple’; *sandaru* /saNdaru/ [sandaru] ‘sandal(s)’ (3, 3 and 4 morae, respectively)

<sup>8</sup> In Japanese phonology, a distinction needs to be made between suprasegmental units, syllable and mora. A word such as *sinbun* ‘newspaper’ consists of two syllables *sin* and *bun*, but a Japanese speaker further subdivides the word into the four units *si*, *n*, *bu*, and *n*, which correspond to the four letters of *kana* used in the written form. A mora in Japanese is a unit that can be represented by one letter of *kana* and functions as a rhythmic unit in the composition of Japanese poems, e.g. *waka* and *haiku*. (Shibatani 1990: 158)

approximately for the same length of time and carries a tone of either high (H) or low (L). The pattern of high and low pitch is arbitrary and unpredictable, so they must be learned for each word. In the mainstream linguistic practice the last mora before a drop in pitch (the change in tone from high to low) is marked, for instance in dictionaries, as the accentual core of the word, e.g. *a'ta' ma* (LHL), but this by all means cannot be compared with the system of stress accent. There are also many Japanese words of which several successive morae carry the high tone, even going over the word boundary, e.g. *ka'rate no sense'i* (LHH H HHHL) 'karate teacher'. The system of pitch accent, that is, a high or low tone carried by each of the rhythmic units of morae, and the opposition of long (two morae) and short (one mora) vowels are foreign to Slovene. These characteristics of Japanese phonology must be somehow modified or dropped when words of Japanese origin are imported for everyday use in Slovene.

## SLOVENE

### 3.1 Gender, number and case

Slovene is a Slavic language and has three genders, i.e. masculine, feminine and neuter, as well as three numbers, i.e. singular, dual and plural. The nouns (and adjectives) are inflected in six cases: nominative, genitive, dative, accusative, instrumental, and locative. Because of case inflection, one word (a semantic entry in a dictionary) may appear in different forms in Slovene sentences. Consider the Slovene translations of the example sentences (4) to (7), from above:

(4) *Včeraj sem kupil kuharsko knjigo.*  
 yesterday AUX buy-PAST cooking book (acc. sg.)  
 "I bought a cookbook yesterday."

(5) *Včeraj sem urejal knjige v knjižnici.*  
 yesterday AUX put in order-PAST books (acc.pl.) in library  
 "I put the books in the library in order yesterday."

(6) *Želim si kuharsko knjigo.*  
 wish-PRES oneself cooking book (acc. sg.)  
 "I want a cookbook."

(7) *Kuharski knjigi sem dodal rože in jih skupaj poslal.*  
 cooking book (dat. sg.) AUX add-PAST flowers and them together send-PAST  
 "I added flowers to the cookbook and sent them."

Such alternation of vowels in a noun (*knjiga – knjige – knjigi – ...*) does not occur in Japanese.



## 3.2 Borrowings

Toporišič (2000), in discussing the origins of various words in Slovene, defines *prevzete besede* ‘assumed words/ borrowings’ as all words taken from other languages into Slovene. Among common words of foreign origin a distinction is made between foreign words (*tujke*) and loanwords (*izposojenke*). While loanwords are those borrowings which have been thoroughly adapted to the Slovene language in their pronunciation, orthography, inflection, syntax and derivation of new words etc., foreign words have not completely adapted, especially in orthography, pronunciation, and inflection. (Toporišič 2000: 131).

About gender assignment of borrowings Priestly (1993) states:

[...] borrowed words normally become either masculine or feminine. The opposition between the two animacy subgenera - animate and inanimate - which is expressed in nouns and their co-referent adjectives and in some pronouns, occurs only in the singular. (Priestly 1993: 399)

Further, Priestly explains the incorporation of borrowings into the contemporary Slovene language as follows:

[...] we find that extremely few borrowed nouns are treated as indeclinables. Normally, if a borrowed noun ends in unstressed *-a*, it is feminine (declined as *lipa*) and otherwise the noun is masculine and declined as *kòt*. Note that nouns ending in *-r* or a vowel add *-j* before non-zero endings, as in *abonmâ*, *abonmâja* ‘subscription’. [...] Note also that virtually no recently borrowed nouns are neuter: hence *nov kino* ‘new cinema’, *nov komite* ‘new committee’, *nov alibi* ‘new alibi’, *nov kanu* ‘new canoe’. (Priestly 1993: 444)

## 4. ASSIGNMENT OF NUMBER AND GENDER

### 4.1. Morphological rules

As we saw earlier, all Japanese nouns end in one of the five vowels, *a*, *e*, *i*, *o*, *u*, or in the syllable-closing, mora-building nasal, *-n*. As for the assignment of the grammatical category of number, it is expected that singular, the unmarked variant, is usually assigned to any noun. The category of gender, as we have already seen, should be decided according to the morphophonological form of the borrowed word. That is, nouns ending in the vowel *-a* are categorized as feminine singular and inflected as feminine nouns. The inflection follows the pattern for many Slovene feminine nouns ending in *-a*, the first feminine declension, e.g. *lipa -e*. Japanese nouns which end in *-n* as well as all other vowels, *e*, *i*, *o* and *u*, should be considered masculine and declined accordingly: those that end in *-n* are to be declined in the pattern used for many Slovene masculine nouns ending in a consonant, the first masculine declension, e.g. *kot -a*; those that end in vowels (and a stressed *-a*) would have extended inflection with the additional *-j*, as in the case of *abonma* above, or *panó -ja* “panneau (sign, notice in French)”.

## 4.2 Semantics and language users' interpretation

Added to the strictly morphophonological rules described in the previous section, some further rules that are otherwise valid for Slovene words also seem to apply to loanwords. Table 1 below shows all possible solutions for the assignment of number and gender to the loanwords of Japanese origin in Slovene. For all possible word endings in Japanese (listed in the far-left column), the solutions are given with examples.

The first solution ① (<1>,<4>,<7>,<10>,<12>,<13>) is given according to the morphological rules for loanwords in Slovene grammar, cited earlier from Priestly (1993) and regularly applied in Slovene.

The second solution ② (<2>,<5>,<8>) takes into account the possible interpretation of the original word ending as plural, especially when the object in question is countable and often used in context as plural, or the word is understood as a plural noun from the beginning. Examples of pluralia tantum in Slovene are: *možgani -ov* 'brain' (m.), *vile vil* 'fork' (f.), *vrata vrat* 'door' (n.).<sup>9</sup> There are also several place names that are plural in Slovene, e.g. *Jesenice Jesenic*. Therefore, it is possible that Slovene speakers adopt a borrowed word as a plural noun when its meaning coincides with some plural connotation of countable objects, or the word is associated with some similar plural noun in Slovene. The plural interpretation may apply for the Japanese word endings *-a*, *-e* and *-i*, neuter, feminine and masculine respectively.

In the column ③ (<3>,<6>,<9>,<11>) are other possible solutions: application of other grammar rules for Slovene words, additional interpretation by language users etc.

Solutions in the table from <1> to <13> are explained below. Actual example sentences are taken from the referential corpus *Fidaplus*. Individual nouns are also verified in at least one of the existing dictionaries.

---

<sup>9</sup> Examples from Toporišič (2000: 271).

Table 1: Gender and number assignments to borrowed nouns of Japanese origin in Slovene

Word ending in Japanese	Possible solutions and examples					
	①		②		③	
	by morphological rules; singular	example	plural interpretation	example	additional rules and interpretations	example
-a	f. sg. <1> [+anim] [-anim]	<i>gejša -e</i> <i>ikebana -e</i>	n. pl. <2>	—	m. sg. [+anim] <3>	<i>jakuza -e</i>
-e	m. sg. (-j) <4>	<i>anime -ja</i>	f. pl. <5>	<i>karaoke</i> <i>karaok</i> <i>šitake šitak</i>	m. sg. <6> [+anim]	<i>kamikaze -a</i>
-i	m. sg. (-j) <7>	<i>cunami -ja</i>	m. pl. <8>	—	-Vi>-Vj m.sg. <9>	<i>samuraj -a</i>
-o	m. sg. (-j) <10>	<i>go -ja</i>			m. sg. <11>	<i>kimono -a</i>
-u	m. sg. (-j) <12>	<i>tofu -ja</i>				
-n	m. sg. <13>	<i>šogun -a</i>				

### **Words ending in -a**

<1> The original Japanese nouns ending in *-a* are interpreted as feminine nouns in the singular. The most typical examples are *gejša*<sup>10</sup> ‘geisha’ and *ikebana*, both of which are already included in the Dictionary of Loanwords (*Slovar tujk*) published in 1968 (hereinafter ST 1968). *Gejša* fits well into Slovene since the word refers to a feminine person.

(8) *Premiera opere o mladi gejši Čo Čo San bo v petek v veliki dvorani SNG.*  
[*Dnevnik* 2004-01-30]

»The first night of the opera about the young geisha Cio-Cio San will be on Friday in the large hall of SNG.«

(9) *Če po ulici postopate dalj časa, lahko srečate tudi gejše ...* [Svet in ljudje 2003-10]  
»If you loaf about in the street for some time, you can meet geishas ...«

Other borrowings of this type with the meaning of masculine, not feminine, persons are classified under <3> below.

<sup>10</sup> Some Japanese consonants in loanwords are usually adapted to Slovene orthography according to their pronunciation: *sh/sy* > *š*, *j/z(y)* > *dž*, *ch/t(y)* > *č*, *ts/t* > *c*, *y/y* > *j* (Japanese spelling in Hepburn system/Japanese spelling in *kunrei* system > Slovene orthography, respectively). The vowel sequences *ei* and *ai* are often spelled as *ej* and *aj* in Slovene. See Mlakar/Ilc (2009: 7-8) for details. This article focusses on gender and number assignment and does not go into detail regarding orthography.

*Ikebana* originally means the art of flower arrangement and the definition in the Slovene dictionary remains the same, but the actual use shows that it is also used to denote individual pieces of arranged flowers or plants, thus used also in plural.

- (10) *Študij mode v Evropi je nadgradil s poglobitvijo v umetnost ikebane.*

[*Mali svet* 2004-12]

“He upgraded the study of fashion in Europe by delving into the art of ikebana.”

- (11) *Izdelovali so različne ikebane in pripravljali novoletne piškote.*

[*Vestnik Murska Sobota* 2006-08-17]

“They made different ikebanas and prepared New Year’s cookies.”

Among recent borrowings of this type are *manga* and *tempura*. Neither of the nouns is yet included in Slovene dictionaries. *Manga* is used in the singular to denote the genre, in the plural to denote individual examples of the publication, similar to the use of *ikebana* above.

- (12) *Japonski strip ali manga ima velikanski vpliv na japonsko bralno kulturo, ...*

[*Mladina* 1993-03-28]

“The Japanese comic strip or manga has great influence on the Japanese readership, ...”

- (13) *Na leto prodajo več kot milijardo in pol izvodov različnih mang.* [*Joker* 2006-08-01]

“They sell more than a billion and a half copies of various mangas.”

The word *tempura*, originally a method of preparing dishes, is used in Slovene in the same meaning but also means a certain kind of flour (example 14 below). Among examples found in the corpus, only one sentence was found in which the word was used in the plural (example 15 below).

- (14) *Repke škampov povajjamo v tempuro, zmešano z vodo, in jih ocvremo.*

[*Svet in ljudje* 2003-10]

“We roll shrimp tails in tempura mixed with water, and fry them.”

- (15) *... saj je dnevno na voljo tako izbor tradicionalnih sušijev kakor sodobnejših suši kreacij, poleg sašimijev, tempur in nekaj izvirnih kreacij, ...* [*Playboy Slovenija* 2003-04]

“... a choice of traditional sushi is available daily, as well as modern sushi creations besides sashimi, tempura and some original creations, ...”

<2> Theoretically, words ending in *-a* may be understood as the plural form of neuter nouns, e.g. in Slovene *besedilo besedila* ‘text, texts’. If the speakers feel comfortable with the solution, especially when the semantics of the word coincides with the plural connotation of countable objects, such an interpretation is possible. However, no such example was found.

<3> This is the interpretation of animate nouns ending in *-a* and denoting persons, usually masculine. An additional Slovene rule applies, i.e. the second masculine inflection for nouns, of the type *vojvoda -e* ‘duke’ (Toporišič 2000: 288). If the borrowed words ending in *-a* mean persons without additional specification for gender, they are systematically used as masculine nouns, e.g. *jakuza* ‘yakuza’, *ninja/nindža*. They are both included in the New Dictionary of Loanwords from 2002 (*Veliki slovar tujk*, hereinafter VST 2002).

(16) *Japonska gangsterska organizacija jakuza služi denar z igrami na srečo*, [2003-04-15]  
“The Japanese gangster organization yakuza earns money by gambling, “

(17) *Leta 1990 je bilo po nekaterih podatkih na Japonskem približno 87.000 poklicnih jakuzov*, [Delo 2000-12-23]  
“In 1990, according to some data, there were about 87,000 professional yakuzas in Japan, “

### **Words ending in -e**

There are many Japanese words ending in *-e* which have been borrowed into Slovene. The assignment of gender and number to these words is diverse. See <4> , <5> and <6> below.

<4> The original Japanese words ending in *-e* are interpreted as masculine nouns of foreign origin in the singular. *Karate -ja* is a good example and has certainly been in use in Slovene since the 1960s, although it was not listed in neither of the dictionaries published in the 60s (*Slovenski pravopis* [Slovene Orthography] hereinafter SP 1962 and ST 1968), first appearing in the Dictionary of the Standard Slovene Language (SSKJ 1975). Of the same type is *sake -ja* (ST 1968). *Karate* denotes a martial art (or sport), *sake* denotes an alcoholic drink. Both are usually used in the singular.

(18) *Trenutno se bolj izpopolnjuje v karateju*, [Nedeljski dnevnik 2000-11-05]  
“At present, he is rather trying to improve at karate.”

(19) *Pokusili boste moj domači sake in ..* [Dnevnik 2002-04-30]  
“You will try my homemade sake, and...”

(20) *Bil je proizvajalec sakeja (japonske nacionalne alkoholne pijače)*, [Mladina 1995-07-29]  
“He was a manufacturer of sake (Japanese national alcoholic drink),”

<5> Words ending in *-e* are interpreted as feminine plural nouns of which the singular form may theoretically be traced back to words ending in *-a*. If the word is actually used also as a singular noun, the singular form is a new derivation in Slovene. *Karaoke* is

defined as a plural noun in dictionaries, *karaoke -ok* (VST 2002 and SP 2007), and means the act of singing in a certain environment, i.e. with accompanying music. In practice, it is also used in the singular, as *karaoka*, a new derivation in Slovene (example 22).

(21) ...*danes na karaokah prepevajo angleški gospodje srednjih let.* [Delo 2002-10-12]  
“Nowadays English gentlemen of middle age are singing in karaoke.”

(22) *Karaoka je še vedno »en vogue«, ...* [Dnevnik 1998-06-17]  
“Karaoke is still in vogue, ...”

There seems to be the same interpretation for nouns denoting an inanimate object, e.g. *šiitake/ šitake* “shiitake” (VST 2002). In this case, the word means a certain kind of mushroom, and in Slovene *goba -e* “mushroom” is also a feminine noun.

(23) *O vzgoji šitak pa je več prispevkov napisal ravno predsednik prej navedenega društva.*  
[Nedeljski dnevnik 1999-12-19]  
“The president of the formerly mentioned association has written many articles on how to grow shiitake.”

(24) *Gobe šiitake k nam prihajajo z daljnega vzhoda,* [Nedeljski dnevnik 2000-01-30]  
“‘Shiitake’ mushrooms come to us from the Far East.”

(25) *Nekaj podobnega je bil rahlo eksotično obarvan krožnik s šitakami (japonskimi gobami),*  
[Delo 2004-02-01]  
“Something similar was the slightly exotic dish with shiitakes (Japanese mushrooms),”

(26) *Potrebujemo (na osebo): 100 g piščančjega fileja, 40 g kumar, 40 g belega misa, 30 g gob šitaki, ...* [Delo 2004-03-30]  
“Ingredients (per person): 100 g chicken filet, 40 g cucumbers, 40 g white miso, 30 g ‘shiitake’ mushrooms, ...”

In the last example, “shiitake” is used as a masculine noun *šitak -i*. The female form in other examples, when pronounced, was perhaps mistaken for masculine, because the last vowel is not stressed (similar to <6> below).

<6> Words ending in *-e* are interpreted as masculine singular nouns when the last vowel is considered as a non-voiced or unaccented *e* after a consonant, as is the case for nouns from Latin or English (e.g. *Wilde Wilda*), though in Japanese all vowels in all syllables are fully voiced. For example, “kamikaze” in Japanese has the accent pattern of LHHH (with no accent core, or with no fall in the pitch within the word) but when borrowed into Slovene, a stress accent is placed on the third syllable, i.e. *kamikáze* (similarly in English /kami'ka:zi/ Oxford Dictionary of English 2002). Therefore, with the additional interpretation of the unaccented final vowel, some

words ending in *-e* are masculine (Toporišič 2000: 282). Semantically, *kamikaze -a* (ST 1968) denotes a way of aircraft fighting, as well as a pilot of such an aircraft.

(27) *Nekateri igralci so namreč kot kamikaze jurišali v nasprotne igralce.*

[*Delo* 2001-01-22]

“For some players rushed at the enemy players like kamikaze.”

(28) *Kasai skače bolj tvegano, pogumen je kot kamikaza.* [*Dnevnik* 1999-01-06]

“Kasai jumps more riskily, he is brave like a ‘kamikaze’.”

The word form in example (28) is a new creation. The original word is taken as a plural feminine form.

(29) *V obdobjih neuspehov je Hillary Clinton moža reševala z odločnostjo kamikaz.*

[*Delo* 1999-07-17]

“In times of failure Hillary Clinton rescued her husband with the decisiveness of a kamikaze.”

(30) *A zdaj se bo moral bojevati s kamikazami, kar zagotovo ne bo preprosto.*

[*Delo* 2002-04-28]

“But now he has to fight with kamikazes, which will surely not be easy.”

### ***Words ending in -i***

<7> Original Japanese words ending in *-i* are interpreted as masculine nouns of foreign origin in the singular. Examples are: *kabuki -ja*, *kaki -ja* (ST 1968) *azuki -ja*, *cunami -ja*, *origami -ja* (ST 2002). They are all regularly declined with the *-j-* extension and used both in the singular and the plural. The recently very popular *suši/sushi* “sushi” should be included in this group, but it is not mentioned in any Slovene dictionary.

(31) *Belo so se šminkali igralci kabukija, gejše in neveste.* [*Delo* 2001-05-21]

“Kabuki actors, geishas and brides painted their faces white.”

(32) *Med njimi so bili tudi igralci gledališča kabuki.* [*Delo* 2000-12-15]

“Among them were actors of the kabuki theater.”

(33) *Poudarek je predvsem na sushiju, po katerem je japonska kuhinja tudi najbolj prepoznavna.* [*Besedilni drobiž* 2006]

“The emphasis is above all on sushi, for which Japanese cooking is also the most known.”

(34) *Čez dan so na številnih stojnicah ponujali japonske dobrote, med katerimi so obiskovalci najpogosteje posegali po sušiju.* [*Delo* 2003-05-18]

“During the day they offered Japanese food on many stalls among which the visitors most frequently reached for sushi.”

<8> Theoretically, words ending in *-i* may be understood as the plural of masculine nouns, e.g. in Slovene *stol stoli* ‘chair, chairs’. However, no such example was found.  
 <9> When original Japanese nouns end in the vowel sequence *-ai*, Slovene orthography rewrites them to the diphthong *-aj*, probably because similar words with such word ending exist in Slovene, e.g. *ležaj -a*, *papagaj -a* etc. The inflection follows the Slovene pattern, and because of *j* at the end of the word the extension rule with an additional *-j-* for loanwords does not apply, but they are regularly inflected as masculine nouns. Examples are: *samuraj -a* (SP 1962) and *bonsaj -a* (VST 2002).

(35) *Mi vsi smo svojevrstni samuraji.* [GEA 2005-08]

“We are all samurais in our own way.”

(36) *Evropejci se za bonsaj ne odločimo vsak dan.* [Nedeljski dnevnik 2000-02-06]

“We Europeans do not decide for bonsai everyday.”

(37) *Zgodovina bonsajev na Japonskem je nekaj stoletij krajša kot na Kitajskem,*

[Nedeljski dnevnik 2000-02-06]

“The history of bonsais in Japan is some centuries shorter than in China.”

### **Words ending in -o**

<10> Original Japanese words ending in *-o* are interpreted as masculine nouns of foreign origin in the singular. If the last syllable carries the stress accent, the word should be declined as *panó panója*, if the accent falls on one of the previous vowels, the word should be declined as *kíno kína*, without the extension with the additional *-j-*. As we saw earlier, most of the Japanese syllables are open and all carry the same phonetic length. The fact is that among the words borrowed into Slovene and ending in the vowel *-o*, many have a long vowel *-oo* at the end of the noun in original Japanese (often written as *ô*, *ō*, or *o:*). Thus, originally, *kendoo*, *budoo*, and so on. These long vowels have been lost mainly due to transcription in European languages, and because the phonological system is different in the recipient language. One of the earliest borrowings, *judo*, as well as *kendo* and *šinto* all lost the long (or stressed) *-o* and are now categorized under 11 below.<sup>11</sup> The only word with this inflection is the one-syllable (and one-mora) word *go -ja* “go (Japanese board game)”.

<11> Original Japanese words ending in *-o*, without any distinction of whether or not the original word ends in a long or a short vowel, are declined after the pattern of Slovene words ending in *-o*, e.g. *besedilo -a*. Examples are: *kimono -a*, *judo -a*, *šinto -a* (SP 1962, ST 1968), *kendo -a* (ST 1968) and others.

(38) *Kimoni so v deželi vzhajajočega sonca vedno manj priljubljeni,* [Delo 2001-05-22]

“In the land of rising sun kimonos are less and less popular.”

<sup>11</sup> Mlakar/Ilc (2009: 9) proposes case inflection of masculine personal names with a long vowel – *ō* with the *-j-* extension: *Tarō Tarōja*.



- (39) *Srajca v stilu japonskega kimona je vedno modna.* [Lisa 2004-09]  
 “A shirt in the style of a Japanese kimono is always fashionable.”
- (40) *Tako je Udo Quellmalz v Atlanti osvojil zlato kolajno v judu.* [Dnevnik 1997-02-10]  
 “In this way, Udo Quellmalz won a gold medal in judo in Atlanta.”
- (41) *sem mu omenil, da bi rad treniral judo.* [Delo 2001-03-02]  
 “I told him that I would like to practice judo.”
- (42) *so vatikanski teologi celo našli skupne točke z budizmom in šintom.* [Mladina 2000]  
 “Vatican theologians even found common points between Buddhism and Shinto.”
- (43) *Danes je kendo tekmovalni šport in deluje po tekmovalnih pravilih.* [Celjan 2001-10]  
 “Today kendo is a competitive sport and functions by competitive rules.”

### **Words ending in -u**

<12> Original Japanese words ending in *-u* are interpreted as masculine nouns of foreign origin in the singular, e.g. *tofú -ja*, *šiátsu -ja*, *zai bátsu -ja* (VST 2002). They are regularly declined with the extension with *-j*, as in *kanu kanuja*, even when the stress accent does not fall on the last syllable.

- (44) *Radič dobro operemo, ga razdelimo na liste in stresemo k tofuju.* [Delo 2004-02-17]  
 “We rinse the chicory well and divide it into small leaves which we add to the tofu.”
- (45) *Torta z brokolijem in tofujem* [Zdravje 2002]  
 “A pie with broccoli and tofu”
- (46) *To je nekakšna kombinacija japonskega šiácuja, kiropraktike, indijske joge in akupresure.* [Delo 2000-01-12]  
 “This is some sort of a combination of Japanese shiatsu, chiropractics, Indian yoga and acupressure.”

### **Words ending in -n**

<13> Original Japanese words ending in *-n* are interpreted as masculine nouns in the singular. This case does not raise any problem since there are many Slovene nouns ending in the consonant *-n* and they are usually masculine, e.g. *balón*, *cigán*, *bazén*, *račún*.<sup>12</sup> Loanwords of this type are *šogun -a* (ST 1968), *futon -a*, *zen -a* (VST 2002) etc.

- (47) *Doba šogunov se je končala 9. novembra 1867.* [Dnevnik 2000-11-09]  
 “The era of shoguns ended on 9 November 1867.”

---

<sup>12</sup> In Slovene, nouns with a non-accented last syllable *-en* are feminine, e.g. *básen basni* “fable”.

(48) *Poznamo tudi alternativne izvedbe vložkov, kot so futoni, slamnjače in rogoznice.*  
[Dolenjski list 1998-05]  
“We also know alternative realizations of cushions like futons, straw and reed mattresses.”

(49) *Po vsem svetu se zenu in zen meditaciji posveča vse večje število tistih, ki želijo...*  
[Mama 2005-03]  
“All over the world there are more and more who devote themselves to Zen and Zen meditation...”

To summarise, the gender and number assignment occurs in the following order: according to the morphological form of the loanword, a gender of either masculine or feminine in singular is assigned. The semantics of animacy plays a big role in this process. If a loanword ends in *-a* or *-e* and means an animate being, particularly human and not specifically feminine, the semantics is stronger than the simple morphophonological rules and therefore the masculine gender is assigned. Phonology is also important in an additional two cases: the vowel sequences *-ai* and *-ei* (orthographic change to *-aj* and *-ej*), and the word ending *-o* for words with two or more syllables (the rule for foreign words does not apply and there is no extension with additional *-j*). These may be understood as ‘naturalization’ processes.

## 5. LIMITS AND FURTHER CREATION

Some of the loanwords are described as “indeclinable (*neskl.*)” in dictionaries. These are cases when the original forms of loanwords are such that inflection according to gender (and number) would make communication ineffective or impossible. For example:

(50) *Predstavniki jamato-e so bili člani šole tosa, [Ft.Z.N.S.H humanistična in družboslovna]*  
“The representatives of yamato-e were members of the Tosa school.”

(51) *Dela ukiyo-e so bila namenjena širokemu krogu porabnikov, [Dnevnik 2002-11-18]*  
“The works of ukiyo-e were meant for a large audience.”

*Jamato-e/yamato-e* and *ukiyo-e* (VST 2002) are Japanese art schools or styles. Though there has been much public interest in these genres of art in Slovenia, these words have not been completely integrated into the Slovene language system, including its orthography. The second part of the words ‘*e*’ means “picture”, a one-syllable word, and this seems too short to undergo case inflection. In indeclinable cases, a similar or general naming of the same object or notion is used before the loanword to help communication:

(52) *Ekstrakt iz gob šitake se je v laboratorijskih poskusih izkazal kot izjemno učinkovito sredstvo. [Delo 2001-05-11]*  
“In the lab experiments an extract from ‘shiitake’ mushrooms proved to be highly effective.”

Many nouns of Japanese origin have further derivations, mainly adjectives and nouns, in accordance with Slovene grammar. For example, the loanwords *kamikaze*, *samuraj*, *judo*, *šogun*, *zen* etc. have well-established derivations:

- (53) , *saj smo se pogosto primorani na las natančno izogniti kakemu kamikaznemu kombiju* [Joker 2006-08-01]  
“We were often forced to somehow escape a kamikaze van by a hairs-breadth.”
- (54) *starodavne borilne tehnike s samurajskim mečem.* [Delo 2004-02-15]  
“Ancient warring techniques with a samurai sword.”
- (55) *Zenovsko-taoistični pogled na svet pa je pravzaprav tisto,* [Večer 1999-05-10]  
“The Zen-Taoistic view of the world is actually that of ...”
- (56) *V sedmem stoletju je bil sumo pod vojskujočim se šogunskim režimom nekaj časa prepovedan* [Gloss Slovenija 2004-07]  
“In the 7<sup>th</sup> century under the warring shogun regime, sumo was prohibited for some time.”
- (57) *Futonska postelja,* [Besedilni drobiz 2006]  
“Futon bed,”
- (58) *Slovenski judoisti so se včeraj v beloruski metropoli udeležili zadnjega turnirja.* [Delo 1999-08-21]  
“The Slovene judoists yesterday attended the last tournament in the Belarus capital.”

Of course, nouns are used in many cases in the same nominative singular form in adjectival usage:

- (59) *Tisto, kar je v kabuki teatru trajanje, je v zahodnem dogajanje,* [Dnevnik 1999-07-30]  
“Whatever is lasting in kabuki theater is action in the West.”
- (60) *Tudi z majico s širokimi kimono rokavi in s prevezovanjem na ramenskem šivu boste zadeli v polno.* [Nova 2003]  
“Even with a shirt with wide kimono sleeves and knots on the shoulders you will shine.”

## 6. CREATIVE COMPETENCE

With the loanwords of Japanese origin in Slovene, the following has been observed:

- A. For the effective use of loanwords in communication, grammatical knowledge of the source language is not required. Such knowledge may even hinder free and creative competence in the recipient language.

B. Foreign words are borrowed into the existing grammatical framework of the recipient language. Users of the recipient language find a proper place for each new word, if and only if they have a semantic reason for borrowing. If the environment in which a word may be received is free of possible misunderstanding and misinterpretation, all the existing rules apply. However, when the environment is ‘crowded’ with similar words and possible miscommunication, some additional mechanism is switched on, e.g. non-inflection, adjectival use of nouns, orthographic change etc., as seen and discussed above.

The most creative, and expressive, assignment of gender and number was observed with the plural noun *karaoke*. As discussed earlier, the possibility of plural assignment to a Japanese loanword of this type may occur if Slovene speakers recognize the denoted object as usually occurring in several pieces, e.g. *šitake*, or if the word is associated with plural proper nouns for places, e.g. *Jesenice*. (cf. Section 4.2, interpretation of ②) It seems that the word ending *-e* in “karaoke” is associated with the latter case, for the connotation of ‘a place in which singing entertainment takes place’. As for proper nouns, the Japanese city of Kobe is sometimes analogically declined as a plural noun:

- (61) *Veliki potresi, film o potresu v Kobah l. 1995* [Dnevnik 2000-11-25]  
“Major Earthquakes, a film about the earthquake in Kobe in 1995.”

## 7. CONCLUSION

Due to the recent development of technology and rapid globalization, there are ever more foreign words and notions accepted in our everyday language use. Even Slovenia and Japan, two nations that are geographically and culturally very distant, are becoming closer. According to Wikipedia, there are more than 150 borrowed words of Japanese origin in Slovene. (Wikipedia: last modified: 7. October 2010) Based on the author’s own judgement, some 60 of them (roughly 40%) are generally known to all or most educated adult Slovene speakers and used regularly in today’s Slovene.

This article does not cover the history of loanwords from Japanese into Slovene. The examples are not exhaustive and no statistical analysis on the use of loanwords has been carried out. Further research may disclose yet new facts and processes of borrowing from Japanese into Slovene.

## Bibliography

Primary sources:

- MLAKAR, Barbara/Iztok ILC (2009) “Predlogi za zapisovanje in pregibanje besed japonskega jezika.” *Azijske in afriške študije XIII-2: Predlogi za zapisovanje in pregibanje besed iz japonščine in kitajščine*. Ljubljana: ZIFF, 3-14.
- PRIESTLY, Tom M. S. (1993) “Slovene.” In: B. Comrie/G. Corbett (eds), *The Slavonic languages*. London: Routledge, 388-451.

SHIBATANI, Masayoshi (1990) "The Japanese Language 8: Phonology." *The Languages of Japan*. Cambridge: Cambridge University Press, 158-184.  
TOPORIŠIČ, Jože (2000) *Slovenska slovnica*. Maribor: Založba Obzorja.

Corpus and references:

DOLESCHAL, Ursula/Anna M. THORNTON (eds) (2000) *Extragrammatical and Marginal Morphology*. München: Lincom Europa.  
SHIGEMORI BUČAR, Chikako (ed.) (2009) *Azijske in afriške študije XIII-2: Predlogi za zapisovanje in pregibanje besed iz japonščine in kitajščine*. Ljubljana: ZIFF.  
SHIGEMORI BUČAR, Chikako (2009) "Surobeniago (Slovenian)." In: S. Kaji, Y. Nakajima, T. Hayashi (eds) *Jiten Sekai no Kotoba 141*. Tokyo: Taishukan, 322-325. [COBISS.SI-ID 39597666]  
ZWICKY, Arnold M./Geofrey K. PULLUM (1987) "Plain morphology and expressive morphology." In: J. Aske et al. (eds), *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistic Society*. Berkeley: Berkeley Linguistic Society, 330-340.

*Slovenski pravopis SAZU*. BAJEC, Anton et al. (1962). Ljubljana: Državna založba Slovenije. [SP 1962]

*Slovar tujk*. VERBINC, France (1968). Ljubljana: Cankarjeva založba. [ST 1968]

*Slovar slovenskega knjižnega jezika*. ([1970, 1975, 1979, 1985, 1991]1995). Ljubljana: Inštitut za slovenski jezik Frana Ramovša ZRC SAZU [SSKJ]

*Veliki slovar tujk*. TAVZES, Miloš (ed) (2002). Ljubljana: Cankarjeva založba. [VST 2002]

*Slovenski pravopis*. TOPORIŠIČ, Jože et al. (eds) (2007). Ljubljana: SAZU/ZRC SAZU. [SP 2007]

*Oxford Dictionary of English* (2003). Oxford: Oxford University Press.

*Fidaplus*. 10 July 2011. <http://www.fidaplus.net/>

*Wikipedia: Seznam slovenskih besed japonskega izvora*. 10 July 2011.

[http://sl.wikipedia.org/wiki/Seznam\\_slovenskih\\_besed\\_japonskega\\_izvora](http://sl.wikipedia.org/wiki/Seznam_slovenskih_besed_japonskega_izvora)

#### **List of abbreviations:**

ACC accusative case  
C consonant  
CONT continuous verb form  
DAT dative case  
f. feminine  
GEN genitive case  
H high tone  
INST instrumental case  
L low tone  
LOC locative case  
m. masculine  
n. neuter  
NOM nominative case

pl.	plural
PRES	present tense
sg.	singular
V	vowel

#### Abstract

### CREATIVE COMPETENCE IN BORROWINGS: WORDS OF JAPANESE ORIGIN IN SLOVENE

Japanese has no category of gender or number. Nouns as well as adjectives have no inflection. The phonological construction of words is syllabic and there is no stress accent but pitch accent. Five vowels have short and long opposition. On the other hand, Slovene has categories of number and gender. Nouns are inflected according to six cases. Phonologically, all nouns carry a stress accent. In this paper, the gender and number assignment to loanwords of Japanese origin in Slovene is discussed on the basis of the morphology and phonology of the source and recipient languages. Assignment rules have certain morphological and semantic foundations. The relative competence of the speakers of the recipient language, Slovene, is particularly noticeable when activating the categories and rules that are specific to their own language. The integration of loanwords may be observed through the additional derivation of adjectives and nouns within the framework of the recipient language.

#### Povzetek

### USTVARJALNA KOMPETENCA V IZPOSOJANJU: BESEDE JAPONSKEGA IZVORA V SLOVENŠČINI

Japonščina nima slovničnih kategorij spola in števila. Samostalniki in pridevniki se ne sklanjajo. Besede so sestavljene z odprtimi zlogi s tonemskim, ne z jakostnim naglasom. Vsi japonski samoglasniki imajo kratko in dolgo različico. Na drugi strani pozna slovenščina kategoriji števila in spola. Oblike samostalnikov se spreminjajo glede na število, spol in sklon. Vsi samostalniki nosijo jakostni naglas.

Pričujoči prispevek obravnava glede na oblikoslovne in fonološke značilnosti obeh jezikov določanje spola in števila pri sprejemanju izposojenk iz japonščine v slovenščino. Pravila pri določanju spola in števila temeljijo na oblikoslovju in semantiki: za število ima prednost nezaznamovana različica, ednina, za spol moški ali ženski spol glede na samoglasnik (le izjemoma soglasnik) zadnjega zloga prvotne japonske besede. Jezikovno ustvarjalna sposobnost govorcev jezika prejemnika, slovenščine, je zlasti opazna, kadar se upoštevajo kategorije in pravila, ki so specifično slovenski. Semantika vršilca oz. udeleženca kot živega bitja ima pomembno vlogo pri določanju spola, ravno tako tudi množinskost samostalnika, če to dovoljuje oz. pričakuje sobesedilo, v katerem se nova beseda uporablja (npr. *karaoke*, *šitake*). Popolno integracijo izposojenk prepoznamo pri izpeljavi pridevnikov in dodatnih samostalnikov, ki so nove tvorbe znotraj jezika prejemnika (npr. *kamikazni kombi*, *samirajski meč*, *judoist*). Najbolj izvirno določanje spola in števila je opazno pri samostalniku *karaoke*.

## NATURALNESS: SOME SYNCHRONIC DATA OF OLD INDIAN

The description of Natural Syntax is here repeated verbatim from Cvetko-Orešnik/Orešnik (ms.).

Natural Syntax is a (developing) deductive linguistic theory that determines the presuppositions on the background of which a (morpho)syntactic state of affairs can be made predictable, and thus synchronically explained. The two basic kinds of presuppositions are what are known as naturalness scales and rules of alignment among corresponding values of any two scales. Every (morpho)syntactic state of affairs is represented by two comparable variants. Natural Syntax contains no generative component.

The basic format of the naturalness scales is  $>\text{nat} (A, B)$ , where A is more natural than B. Two expanded scales are allowed, viz.  $>\text{nat} (A + B, B)$  and  $>\text{nat} (A, A + B)$ ; they are valid if the corresponding scale of the format  $>\text{nat} (A, B)$  is valid. Exemplification below.

I proceed to list the criteria with which Natural Syntax substantiates naturalness scales:

- (a) The speaker/hearer parameter. In the scale,  $>\text{nat} (A, B)$ , value A is natural for the speaker (and unnatural for the hearer); value B is unnatural for the speaker (and natural for the hearer). The basic naturalness scale is,  $>\text{nat}$  (favourable for the speaker, favourable for the hearer). This view of naturalness is commonplace in linguistics (Havers 1931: 171), under the names of tendency to economize (utilized first of all by the speaker) and tendency to be accurate (mainly in the hearer's interest).

We follow Mayerthaler (1981: 13 ff.) in assuming that the speaker is the centre of communication, therefore most properties of the speaker are natural; for instance being the first person and/or the subject and/or +human and/or +masculine (!) and/or +singular and/or +definite and/or +referential, etc.

What is favourable for the hearer may be less natural for the speaker. This is a pivotal point in Natural Syntax and will be maintained until some good counter-example nullifies it. By way of illustration it can be pointed out that producing a longish noun phrase may be "tiresome" for the speaker (= less natural for him), but may ease the hearer's decoding process considerably (= be more natural for the hearer).

---

\* *Author's address:* Filozofska fakulteta, Oddelek za primerjalno in splošno jezikoslovje, Aškerčeva 2, Ljubljana, Slovenia. Email: [cvetko@zrc-sazu.si](mailto:cvetko@zrc-sazu.si)

- (b) The principle of least effort (Havers 1931: 171). What conforms better to this principle is more natural for the speaker. What is cognitively simple (for the speaker) is easy to produce, easy to retrieve from memory, etc.
- (c) Degree of integration into the construction. What is better integrated into its construction is more natural for the speaker.

As a rule of thumb, what is located at the margin of a construction is less natural (for the speaker) than what is placed inside a construction.

- (d) Frequency. What is more frequent tokenwise is more natural for the speaker. What is cognitively simpler (for the speaker) is used more. (However, the reverse does not obtain: what is natural for the speaker is not necessarily more frequent.)
- (e) Small vs. large class. The use of (a unit pertaining to) a small class is more natural for the speaker than the use of (a unit pertaining to) a large class. During speech small classes are easier for the speaker to choose from than are large classes. (This is frequency typewise.)
- (f) The process criterion. Any process is natural. Examples of processes: movement, agreement.
- (g) Acceptable vs. non-acceptable use. What is acceptable is more natural for the speaker than what is not acceptable. The very reason for the acceptability of a syntactic unit is its greater naturalness for the speaker with respect to any corresponding non-acceptable unit.
- (h) What is more widespread in the languages of the world is more natural for the speaker (the typological criterion). What is cognitively simpler (for the speaker) is realized in more languages.

We have been applying the above criteria (henceforth also called axioms) (a–h) to the language material covering several languages and miscellaneous (morpho-)syntactic states of affairs. Throughout our work, the criteria have compelled us, time and again, to reject certain solutions and to give precedence to other solutions. Given this encouraging experience, we will preserve the present list (a–h) until some convincing and irreparable counter-example casts doubt upon our axioms. The occurrence of such an event is in the overriding interest of Natural Syntax anyway. The only realistic aim of deductive theories is that they are eventually disproved. We are afraid that any improvement of the axioms would lead to a reduction of the chances for the desirable definitive outcome.

The above criteria of naturalness are utilized to support our naturalness scales. Normally it suffices to substantiate any scale with one criterion, which backs up either value A or value B of the scale; the non-supported value is allotted the only remaining position in the scale. Of course, a scale may be supported with more than one criterion. Any clash among the criteria applied to a scale is to be handled with constraints on the combinations of criteria. So far only a few constraints have been formulated; we have not yet encountered much useable crucial language data.



The naturalness scales are an essential part of what are known as deductions, in which Natural Syntax expresses its predictions about the state of affairs in language data. An example of a deduction:

English. The numerical indication of frequency normally consists of a cardinal number followed by the word *times* (e.g., *four times*) except that there are one-word expressions available for the lowest numbers: *once*, *twice* and archaic *thrice* (Collins Cobuild 1990: 270-71).

The two variants: the type *once* and the type *four times*.

1. The assumptions of Natural Syntax:

1.1. >nat (type *once*, type *four times*)

I.e., the type *once* is more natural than the type *four times*. – According to the criterion of least effort, item (b) in the list of axioms.

1.2. >nat (low, non-low) / number

I.e., any low number is more natural than any non-low number (Mayerthaler 1981: 15). – Low numbers are more easily accessible to the speaker. According to the speaker/hearer criterion, item (a) in the list of axioms.

2. The rules of parallel alignment of corresponding values:

2.1. value A tends to associate with value C,

2.2. value B tends to associate with value D. See Note 4.1 below.

3. The consequences:

If the language distinguishes between low and non-low numbers in numerical indications of frequency such that one kind of number uses the pattern *four times* and the other kind of number uses the pattern *once*, it is the low numbers that tend to use the pattern *once*, and it is the non-low numbers that tend to use the pattern *four times*.

Q.E.D. (The reverse situation is not expected.)

4. Notes

4.1. Value A of scale 1.1 (= the type *once*) tends to combine with value C of scale 1.2 (= low number). Value B of scale 1.1 (= the type *four times*) tends to combine with value D of scale 1.2 (= non-low number). Similarly in the remaining deductions, with the proviso that the alignment (unlike here) is chiasitic in most cases. Chiasitic alignment is explained below.

4.2. Natural Syntax cannot predict the cut-off point between low and non-low numerals.

4.3. Henning Andersen (p.c.) has pointed out to us that there is a parallel system covering numerical indications of frequency, *one additional time*, *two/three/four additional times*, etc., which does not make use of the dichotomy treated in the above deduction. Donald Reindl (p.c.) has added *one more time*, *two/three/four more times*, etc.

This deduction maintains that the state of affairs cannot be the reverse; i.e., that numerals above *two* (or *three*) would be one-word formations and that the numerals under *three* (or *four*) would be two-word formations. All predictions of

our Natural Syntax are restricted to such modest claims about the unlikelihood of the reverse situation.

In every deduction, the rules of alignment play a prominent role; compare item 2 in the above deduction. The alignment rules regulate the combinations of corresponding values of the two naturalness scales mentioned in the deduction.

The alignment can be parallel or chiasitic. Suppose that the two scales are  $>nat$  (A, B) and  $>nat$  (C, D). Parallel alignment pairs value A with value C, and value B with value D. Chiasitic alignment pairs A with D, and B with C.

A paramount question is when the alignment is parallel and when chiasitic. Parallel alignment is the default case. Chiasitic alignment is necessary whenever a given deduction is limited to the language data obtaining within an "unnatural environment". This is defined as value B of the scale  $>nat$  (A, B).

An example. In the scale  $>nat$  (main, dependent) / clause, the value "dependent clause" is an unnatural environment. This means: all deductions whose language data lie within the environment "dependent clause" require the implementation of chiasitic alignment.

Chiasitic alignment is prohibited when a naturalness scale is substantiated with an axiom. If, however, an axiom is engaged as one of the scales in a deduction, it obeys the usual distribution of the alignment rules. While Natural Syntax in principle does not deal with semantic phenomena, it does happen in rare cases that semantics is involved in a deduction. Our experience so far suggests that semantic phenomena are never the cause of chiasitic alignment. See deduction (3) below.

The insistence of Natural Syntax on the distinction between parallel and chiasitic alignments stems indirectly from the work of Henning Andersen within the markedness theory. Andersen observes situations such as the following in all human semiotic systems: on an everyday occasion casual wear is unmarked, formal wear marked; on a festive occasion it is the formal wear that is unmarked, whereas casual wear is marked. See Andersen (1972: 45, esp. fn. 23). This example expressed with our scales: (i)  $>nat$  (casual, formal) / wear, (ii)  $>nat$  (-, +) / marked. A third scale as the source of the environment of the deduction:  $>nat$  (everyday, festive) / occasion. If the environment is "every-day occasion", the alignment within (i-ii) is parallel; if the environment is "festive occasion", the alignment within (i-ii) is chiasitic.

After this summary of the basic tenets of Natural Syntax I proceed to the subject-matter of the present paper. It is based on the synchronic morphosyntactic and morphological data assembled in Jamison (1983), a significant monograph on the Rig Veda and Atharva Veda *-áya-* verbs. A few introductory remarks are in order:

In the deductions adduced below unnatural environments are implemented where appropriate, namely "derivative", "causative" and "transitivity". The details:

The basic naturalness scale covering any simplexes and their derivatives is,  $>nat$  (simplex, derivative), i.e., a simplex is more natural than a corresponding derivative. The scale is supported with the circumstance that the internal structure of the derivative is usually more complicated than the internal structure of the corresponding simplex. (Back-formation and most cases of conversion can be exceptional in this

respect.) Therefore the simplex is more natural according to the criterion of least effort, item (b) in the list of axioms. It follows that derivatives constitute an unnatural environment of the deductions that operate exclusively within derivatives. Consequently such deductions call for the chiastic alignment of the values contained in the naturalness scales adduced in those deductions.

On the average causative verbs are structurally (and according to the number of core participants) the most complicated verbs, and as such unnatural according to the criterion of least effort, item (b) in the list of axioms. It follows that causatives constitute an unnatural environment of the deductions that operate exclusively within causatives. Consequently such deductions call for the chiastic alignment of the values contained in the naturalness scales adduced in those deductions.

The basic naturalness scale covering transitive and intransitive verbs is, >nat (intransitive, transitive) / verb, i.e., an intransitive verb is more natural than a transitive verb. Intransitive verbs take fewer core participants than transitive verbs, therefore intransitive verbs are natural according to the criterion of least effort, item (b) in the list of axioms. It follows that transitivity constitutes an unnatural environment of the deductions that operate exclusively within transitivity. Consequently such deductions call for the chiastic alignment of the values contained in the naturalness scales adduced in those deductions.

I proceed to present the deductions:

(1) Rig Veda and Atharva Veda. Causative *-áya-* verbs. The root is monosyllabic. When the root vowel is *a*, it is long in open syllables, short in closed syllables. Examples: *drāv-áyati* 'makes run' as against *krañd-áyati* 'makes roar' (Jamison 1983: 105 ff.).

The two variants: open and closed syllables of the root. – The deduction proceeds in the unnatural environments "derivative".

1. The assumptions of Natural Syntax:

1.1. >nat (open, closed) / root syllable

I.e., an open syllable is more natural than a closed syllable. – In some languages open syllables are almost the only type of syllables allowed whereas in no languages do closed syllables prevail to the same extent. Therefore open syllables are more natural according to the typological criterion, item (h) in the list of axioms.

1.2. >nat (short, long) / root vowel *a*

I.e., a short vowel is more natural than a long vowel. – Long vowels are of longer duration, hence require more effort from the speaker than short vowels. According to the criterion of least effort, item (b) in the list of axioms.

2. The rules of chiastic alignment:

2.1. value A tends to associate with value D,

2.2. value B tends to associate with value C.

3. The consequences:

If the language distinguishes (within causative *-áya-* verbs whose root vowel is *a* or *ā*)

between open and closed root syllables such that one option requires a short root vowel and the other option requires a long root vowel, then it is the open root syllable that tends to require a long root vowel, and it the closed root syllable that tends to require a short root vowel. Q.E.D. (The reverse situation is not expected.)

4. Note. With 21 verbs the root vowel *a* obtains instead of the expected root vowel *ā* (Jamison 1983: 206, 209).

(2) Rig Veda and Atharva Veda. *-áya-* verbs. There are 31 intransitive verbs and 175 transitive verbs (Jamison 1983: 41, 78–104, 105–68). These statistics are representative of a more general situation, viz., of the relative type frequency of transitive and intransitive verbs. The deduction treats this broader case.

The two variants: transitive and intransitive verbs.

1. The assumptions of Natural Syntax:

1.1. >nat (intransitive, transitive) / verb

I.e., an intransitive verb is more natural than a transitive verb. – On the average, an intransitive verb has fewer core participants than a transitive verb. Therefore the intransitive verb is natural according to the criterion of least effort, item (b) in the list of axioms.

1.2. >nat (small, large) / class

I.e., a small class is more natural than a large class. – This is the very criterion of small vs. large class, item (e) in the list of axioms.

2. The rules of parallel alignment:

2.1. value A tends to associate with value C,

2.2. value B tends to associate with value D.

3. The consequences:

If the language distinguishes between transitive and intransitive verbs such that one class is small and the other class is large, then it is the intransitive verbs that tend to constitute a small class, and it is the transitive verbs that tend to constitute a large class.

Q.E.D. (The reverse situation is not expected.)

4. Notes

4.1. Whereas most of the numerical data of Jamison (1983) are limited to relatively low numbers and small numerical differences (this being a consequence of the philological situation in the texts studied), the size of the two classes dealt with in this deduction makes handling in Natural Syntax a realistic undertaking.

4.2. The predictions stated in item 3 above are meant to obtain in all languages that make a distinction between transitive and intransitive verbs.

(3) Rig Veda and Atharva Veda. "There are no causatives to trans. verbs in early Vedic" (Jamison 1983: 186).

The two variants: +/-acceptable causative verb. – The deduction proceeds in the unnatural environment "causative".

1. The assumptions of Natural Syntax:

1.1. >nat (intransitive, transitive) / base verb

I.e., an intransitive verb is more natural than a transitive verb. – Intransitive verbs are accompanied by fewer core arguments than transitive verbs. Therefore intransitive verbs are natural according to the criterion of least effort, item (b) in the list of axioms.

1.2. >nat (-, +) / corresponding causative verb

I.e., a non-causative verb is more natural than a causative verb. – Causative verbs are structurally the most complicated type of verbs, therefore unnatural according to the criterion of least effort, item (b) in the list of axioms.

2. The rules of chiasitic alignment:

2.1. value A tends to associate with value D,

2.2. value B tends to associate with value C.

3. The consequences:

If the language distinguishes between transitive and intransitive base verbs such that one kind is accompanied by corresponding causatives and the other kind is not accompanied by corresponding causatives, then it is the intransitive verbs that tend to be accompanied by corresponding causatives, and it is the transitive verbs that tend not to be accompanied by corresponding causatives. Q.E.D. (The reverse situation is not expected.)

(4) Post-Rigvedic double transitives (the early construction) contain the obligatory accusative of the base-verb subject and the optional accusative of the base-verb object (Jamison 1983: 35).

The two variants: the subject and the object of the base verb. – The deduction proceeds in the unnatural environment "transitivity".

1. The assumptions of Natural Syntax:

1.1. >nat (subject, object) / of base verb

I.e., the subject is more natural than the object. – Subjects are more frequent textwise than objects, for many clauses lack an object. According to the frequency criterion, item (d) in the list of axioms.

1.2. >nat (-, +) / accusative of derived verb

I.e., a lacking accusative is more natural than a present accusative. – According to the criterion of least effort, item (b) in the list of axioms.

A special case of 1.2:

1.2.1. >nat (+/-, +) / accusative of derived verb

Scale 1.2.1 assumes the permitted expanded format >nat (A + B, B) and is automatically valid because the corresponding basic scale 1.2 has been substantiated.

2. The rules of chiasitic alignment:

- 2.1. value A tends to associate with value D,
- 2.2. value B tends to associate with value C.
3. The consequences:

If the language distinguishes between the subject and the object of the base verb such that one clause element corresponds to the obligatory accusative of the derived verb and the other clause element corresponds to the optional accusative of the derived verb, then it is the subject of the base verb that tends to correspond to the obligatory accusative of the derived verb, and it is the object of the base verb that tends to correspond to the optional accusative of the derived verb. Q.E.D. (The reverse situation is not expected.)

(5) Post-Vedic double transitives (the late construction) contain the instrumental of the base-verb subject and the accusative of the base-verb object (Jamison 1983: 189).

The two variants: the subject and the object of the base verb. – The deduction proceeds in the unnatural environment "transitivity".

1. The assumptions of Natural Syntax:
- 1.1. >nat (subject, object) / of base verb

I.e., the subject is more natural than the object. – Subjects are more frequent textwise than objects, for many clauses lack an object. According to the frequency criterion, item (d) in the list of axioms.

- 1.2. >nat (accusative, instrumental) / of derived verb

I.e., the accusative is more natural than the instrumental. – Accusatives are more frequent textwise than instrumentals. According to the frequency criterion, item (d) in the list of axioms.

2. The rules of chiasmic alignment:

- 2.1. value A tends to associate with value D,
- 2.2. value B tends to associate with value C.

3. The consequences:

If the language distinguishes between the subject and the object of the base verb such that one clause element corresponds to the accusative of the derived verb and the other clause element corresponds to the instrumental of the derived verb, then it is the subject of the base verb that tends to correspond to the instrumental of the derived verb, and it is the object of the base verb that tends to correspond to the accusative of the derived verb. Q.E.D. (The reverse situation is not expected.)

4. Note. The optionality of the accusative is treated in deduction (6).

(6) Post-Vedic double transitives (the late construction) contain the instrumental of the base-verb subject and the optional accusative of the base-verb object (Jamison 1983: 189). This matter is treated in deduction (5). The present deduction deals with the optionality of the derived accusative of the base-verb object.

The two variants: +/-accusative of the derived verb. – The deduction proceeds in the unnatural environment "transitivity".

1. The assumptions of Natural Syntax:

1.1. >nat (-, +) / accusative of derived verb

I.e., a lacking accusative is more natural than a present accusative. – According to the criterion of least effort, item (b) in the list of axioms.

1.2. >nat (more, less) / frequent tokenwise

I.e., what is more frequent is more natural than what is less frequent. – This is the very frequency criterion, item (d) in the list of axioms.

2. The rules of chiasitic alignment:

2.1. value A tends to associate with value D,

2.2. value B tends to associate with value C.

3. The consequences:

If the language distinguishes between the realized and the non-realized accusative such that one option is more frequent and the other option is less frequent, then it is the realized accusative that tends to be more frequent, and it is the non-realized accusative that tends to be less frequent. Q.E.D. (The reverse situation is not expected.)

4. Note. Jamison (1983: 189) adduces only one example of lacking derived accusative, and several examples of realized derived accusative (see also her p. 36).

## CONCLUSION

It can be seen that Natural Syntax can predict the Old Indian data treated in the above deductions, in the sense that the predictions do obtain if the presuppositions mentioned in each deduction are accepted. To be sure, Natural Syntax has something to say only about those areas of any language that display variants that can be compared to each other. With this limitation, Natural Syntax has successfully dealt with a variegated language material from many languages.

Natural Syntax treats a morphological case in deduction (1) above, which fact (together with some other evidence) enhances the likelihood that Natural Syntax can be implemented even outside syntax, at least to some extent.

## References

ANDERSEN, Henning (1972) "Diphthongization." *Language* 48, 11–50.

*Collins Cobuild English Grammar*. London: HarperCollins, 1990.

CVETKO-OREŠNIK, Varja/Janez Orešnik (Ms.) *Natural Syntax of Slovenian: The complex sentence*.

HABERS, Wilhelm (1931) *Handbuch der erklärenden Syntax*. Heidelberg: Winter.

JAMISON, Stephanie W. (1983) *Function and form in the -áya-formations of the Rig Veda and Atharva Veda*. Göttingen: Vandenhoeck/Ruprecht.

MAYERHALER, Willi (1981) *Morphologische Natürlichkeit*. Wiesbaden: Athenaion.

Abstract  
NATURALNESS: SOME SYNCHRONIC DATA OF OLD INDIAN

The framework of the paper is Natural Syntax initiated by Janez Orešnik in the tradition of (morphological) naturalness as established by Wolfgang U. Dressler and †Willi Mayerthaler. The basic tenets of Natural Syntax are described at the beginning of the paper.

Natural Syntax is here applied to aspects of Old Indian synchronic verbal morphology and verbal morphosyntax:

- (1) Causative *-áya-* verbs in Rig Veda and Atharva Veda. The root vowel *a* is short in closed syllables and long in open syllables.
- (2) Rig Veda and Atharva Veda contain 31 intransitive and 175 transitive *-áya-* verbs.
- (3) Rig Veda and Atharva Veda do not contain any causative verbs to the basic transitive verbs.
- (4) Post-Vedic double transitive verbs (the early construction) contain an obligatory accusative corresponding to the subject of the base verb, and an optional accusative corresponding to the object of the base verb.
- (5) Post-Vedic double transitive verbs (the late construction) contain an obligatory instrumental corresponding to the subject of the base verb, and an accusative corresponding to the object of the base verb.
- (6) The accusative mentioned in (5) is not obligatory.

The data have been taken from Jamison (1983), a monograph about Old Indian *-áya-* verbs.

Povzetek  
NARAVNO JEZIKOSLOVJE – NEKAJ SINHRONIH PODATKOV IZ STARE INDIJŠČINE

Okvir tega članka je naravna skladnja, kot jo je začel Janez Orešnik v tradiciji (morfološke) naravnosti, utemeljene v delih Wolfganga U. Dresslerja in †Willija Mayerthalerja. Osnovna načela naravne skladnje so opisana na začetku članka.

Tu so obravnavani nekateri vidiki staroindijske sinhronne morfonologije in morfosintakse glagola:

- (1) Kavzativni *-aya-* glagoli v Rg Vedi in Atharva Vedi. Korenski samoglasnik *a* je praviloma/večinoma kratek v zaprtih zlogih in dolg v odprtih zlogih.
- (2) Rg Veda in Atharva Veda izkazujeta 31 neprehodnih in 175 prehodnih *-aya-* glagolov.
- (3) V Rg Vedi in Atharva Vedi ni nobenega kavzativnega glagola iz bolj osnovnih prehodnih glagolov.
- (4) Po vedskem obdobju vsebujejo dvojno prehodni glagoli (to je zgodnja zgradba) obvezen akuzativ, ki je soodnosen z osebkom osnovnega glagola, in neobvezen akuzativ, ki je soodnosen s predmetom osnovnega glagola.
- (5) Po vedskem obdobju vsebujejo dvojno prehodni glagoli (to je pozna zgradba) obvezen instrumental, ki je soodnosen z osebkom osnovnega glagola, in akuzativ, ki je soodnosen s predmetom osnovnega glagola.
- (6) Akuzativ, omenjen v točki 5, ni obvezen.

Gradivski podatki so iz monografije Jamison (1983), posvečene staroindijskim *-aya-* glagolom.



## TESTING THE SUFFIXOIDIZATION OF GERMAN *-MANN* 'MAN'

### 1. INTRODUCTION

Linguists have argued the usefulness of the concept *suffixoid* for several decades now (Elsen 2009, Leuschner and Wante 2009, Lightfoot 2000, 2005, Stevens 2000, 2010). Recently, Stevens has proposed criteria to classify this phenomenon (2010). If such parameters for classification prove useful, there is a stronger argument in favor of utilizing the linguistic category of *suffixoid*. This study uses German *Mann/-mann* 'man/person', relegated to the status of *suffixoid/Halbsuffix* (Duden 1995, Lightfoot 2006), as an object of investigation to test Stevens' proposed parameters. Discussion then follows on the advantages and challenges related to the given parameters, and future work is suggested.

### 2. THE CATEGORY OF SUFFIXOID

The linguistic notion of *suffixoid* refers to a morphosyntactic category situated developmentally between a full word and a traditional derivational suffix. The literature also sometimes refers to this construct as a *semi-suffix* or in German as a *Halbsuffix*. Similar terminology is likewise linked to affixes (*affixoid*) and prefixes (*prefixoid*), but not with confixes (e.g., German *bio-*, which is a bound form) (Stevens 2010: 79), nor with infixes (e.g., Latin *rupit* 'broke' with its nasal morphology inserted root-interior which yields *rumpit* 'breaks') (Matthews 1991:132). Most scholarly work on *suffixoids* and the like is in Germanic linguistics, though Slavic scholars have given attention in this area at least in the 1970s and 1980s (Lightfoot 2000: 56). It deserves mention that at least three of the standard references for the German language, one an etymological resource (Kluge 1999), one a general grammar work (Duden 1995), and one an introduction to German word-formation (Erben 2000) have all embraced the notion of *affixoids* as accepted terminology.<sup>1</sup>

Simply put, as a linguistic item between the categories of root and suffix, a *suffixoid* is best viewed from an historical perspective as a kind of example of *heterosemy*. That is, at least two morphosyntactically variant linguistic items are synchronically on hand that stem from the same etymon. And specifically, one variant must be an unbound form, and another variant shares the same form, but is bound, productive, and has a more generalized meaning. A typical example in the literature is

---

\* *Author's address:* Department of Modern Languages and Classics, 224 B.B. Comer, Box 870246, Tuscaloosa, AL 35487-0246, USA. Email: dlightfo@as.ua.edu

<sup>1</sup> For counter opinions, note the discussion in Elsen (2009) and Stevens (2010), and see also what is considered the seminal contra position in Schmidt (1987). More recent rejection of the usefulness of the term *affixoid* is in Booij (2009) and Kastovksy (2009). This paper does not focus on that general debate per se.

German *Werk/-werk* ‘e.g. work, station’ / ‘e.g. *collective*, grouping of a whole whose individual parts are not readily or foremost perceived’ (Erben 2000: 139, Henzen 1947: 192-193, Lightfoot 2005: 589-590, Nübling et al 2008: 77-79, Stevens 2010: 76-77). During the Middle High German (MHG) period (about 1050-1350 A.D.), for example, we witness an orthographical variant of *Werk*, namely *-werc*, where the crux of the semantics lies on the second compound constituent, quite the norm for Germanic root plus root constructions, thus signaling

(1) *vor-werc* ‘outerworks or projection from a fortification’.

Not only does the root exist in compounds at this time, but the genetically related simplex form continued to exist as it has throughout the ongoing High German periodizations.

(2) Old High German (OHG, circa 750-1050 A.D.) *werc(h)* ‘work, handmade construction from simple material, etc.’

(3) MHG *werc* ‘work, business, handmade material’.

In the same MHG period where we find (1) and (3), Henzen notes a series of constructions with the *-werc* form as a secondary semantic constituent bearing collective meaning (1947: 194), for example

(4) *scuohwerc* ‘footwear’, *vleischwerc* ‘meatcuts of a butcher’, *dincwerc* ‘(legal) court (related matters)’.

Thus the formally similar, genetically related items are seen synchronically with a free morphosyntactic manifestation (a root), and at a middle stage displaying semantics associated with a root, but bound formally (a root constituent of a compound), and at a stage bound formally and having lost full semantic gravity associated with its root cognate. These are affixoidal characteristics. Remarkably, though the diachronic picture tends to often yield roots continuing to exist alongside compounding and affixizing formally similar but semantically generalizing cognates, the scholarly discussion has been and remains rather challenged to explore and systematically describe this fairly opaque morphosyntactic transition.<sup>2</sup>

### 3. CRITERIA FOR AFFIXOIDS AS PROPOSED BY STEVENS (2010)

Though scholars have remarked upon the characteristics of affixoids with relative vigor in particular from the 1970s and 1980s onward (Lightfoot 2000: 56, Schmidt

---

<sup>2</sup> Kastovsky goes so far as to suggest that “...[w]e thus have to accept this indeterminacy, which reflects the diachronic development from lexeme to affix” (2009: 328).

1987), Stevens (2000: 58-59) sought to synthesize these notions in diagnostic fashion. Thereafter, interestingly enough in a volume dedicated to grammaticalization phenomena in German (Leuschner et al 2010), Stevens continues calibrations to assist in identifying affixoids, indicating that his criteria are gleaned for the most part from collegial consensus (2010: 72-73). He arrives at the following compilation for affixoids, which is almost immediately followed by an addendum (inserted in c. below in italics) (2010: 74).

Tests for Affixoids:

- a. Affixoids in spoken languages are not just serial but usually very productive. They are the basis for new formations.
- b. Affixoids exist alongside a formally identical, and usually free, “parent” morph. This means there are two linguistic items identical in form and one is derived from the other.
- c. The meaning of the affixoid is more generalized and abstract than the formally identical parent. *A generalized meaning includes metaphorical usage when there is no such usage of the free etymon.* In contrast with some affixes, the affixoid does not express grammatical relations (like person, case, tense, etc.).
- d. There has been a shift of meaning in the relationship between the two parts of the word, so that the first, or other, component determines the basic meaning. That is, the syntactic relationship between the two morphemes is not like that of a subordinating compound. The relationship between the parts of the word are also not like other compounds (copulative, exocentric, etc.). Related to this is the inability to paraphrase the formation of the affixoid, e.g. German *Riesenkraach* ‘(a) very loud noise’ is not equal to \**Kraach eines Riesen* ‘(the) sound of a giant’ or \**Kraach, wie ihn ein Riese macht* ‘(the) sound a giant makes’.
- e. The affixoid must be in competition with or in complementary distribution with affixes.

(slightly adapted from Stevens 2010: 72-74, translations from Stevens 2000: 59)

The main additions in the more recent work building on Stevens (2000) are point e., that is, the affixoid is in a relation of sorts with a system of affixes, and the aforementioned inclusion in c. regarding possible metaphorical extension of the bound form.

These tests appear sensible from a theoretical standpoint. As noted above, they are generally consensual. Let us now put them into practice, taking German *Mann/-mann* as the principle example. Do the above criterion perform well in categorizing real language?

#### 4. APPLYING THE TESTS TO GERMAN *MANN/-MANN*

This study utilizes German *Mann/-mann* ‘man, person’ as a primary example due to its recognition in a standard reference as a suffixoid (*Duden* 1995), which is bolstered due to comparative developments in other Germanic languages (e.g. see Booij 2009: 208-209 for Dutch, and consider an example like English *sportsmanship*,

arguably fully “bleached” of the cognate meaning ‘man’), and a final reason for selection is the form’s high degrees of usage and heterosemy. The diachronic record for *-mann* is also examined below, as the form evidences a long history of developments and the semantics involved are complicated.

To satisfy (a) above, the bound form *-mann* must demonstrate a high degree of productivity. As the *Duden* (1995) pertains to Modern German, it is in this period where certain examples are initially sought. An online lemma search of the mainstream newsmagazine *Spiegel’s* archive (containing searchable issues since 1947) yields approximately 200 usages of *Sportsmann* ‘sportsman’, 75 of *Forstmann* ‘forester’, and 23 of *Weltrekordmann* ‘world record holder’. These three types were searched because they are the complete set of examples noted in the *Duden* (1995: 508). The relatively high frequency of tokens in the newsmagazine genre alone attests to a certain threshold of productivity to satisfy condition (a). These three types also appear in Mater’s (1967) backwards dictionary of German, and 207 other such types further attest to the robust productivity of *-mann* synchronically (as opposed to 71 types for *-frau*).

Criterion (b) is also thus met, as it is clear the very commonly used noun *Mann* (an unbound root) co-occurs in Modern German along with bound examples such as those noted immediately above (Mater 1967). That the bound form derives from the unbound similar form cannot be disputed, as noted in the standard references and other works such as Erben (2000: 38) and Henzen (1947: 195).

Criterion (c) potentially presents something of an issue, etymologically speaking. The historical record indicates that reconstructed Germanic *\*mann* ‘man, person’ (*Duden* 1995, *Kluge* 1999) was not limited to signaling male human beings, and this would lend little to no room for generalization of meaning from MALE > ANY PERSON in the affixoid. As in *Kluge* (1999), though, we see that the main meaning for *mann* in the various periods leans predominantly toward ‘man’ (i.e. adult male), so it could be argued that *mann* historically foremost meant ‘male’. Further evidence for such a consideration of maleness associated with *mann* comes from the construal by modern speakers of the cognate indefinite pronoun *man* ‘one, you’, which is supposed to denote anyone, regardless of sex. Speakers can perceive the usage as awkward in utterances such as the following:

(5) *Man wird Mutter* ‘One becomes a mother’.

or

(6) *Man geht auf das Damenklo* ‘One goes to the ladies’ W.C.’

In some instances, a “female” pronominal variant may serve as a substitute for the prescribed pronoun to ostensibly balance out the perceived linguistic unfairness.<sup>3</sup>

---

<sup>3</sup> Personal electronic communication with a female colleague also displays usage of *frau* functioning as an indefinite pronoun in non-gender related domains. Further discussion in this area is required to determine if and to what degree folk etymology is involved here or not.

A simple Google search reveals, however, that online usage is oriented typically around domains that are female-focused, as in (5) and (6) above. Example (7) comes from a source about women's fashion (mypashion.com), and the normally unwarranted orthographic capitalization of the nascent indefinite pronoun illustrates the novelty and biological gender meaning encroaching on this construction.

(7) [D]as muss Frau einfach lieben. 'One (female) has to simply love it/that.'

Based on (5)-(7), I argue that since a substantial segment of modern speakers must construe free forms of *man* as at least sometimes signaling something like 'one (who is a male)', a modern perception of the bound form *-mann* generalized to mean 'anyone (regardless of gender)' fulfills point (c) on generalization/abstraction of meaning. Such examples are rare in the modern *Spiegel* corpus. One must search back to the summer of 1989 to find a lemma that is semantically ambiguous in terms of the signified's sex. The context is that of couples who have lived together before getting married, and what the supposed advantages of such previous experience are thought to be.

(8) *Intime Kenntnis des Partners, ...was dem Sportsmann zum Vorteil gereicht...* 'intimate knowledge of one's partner, which passes as an advantage for the athlete (hunter)' (*Spiegel* July 3, 1989)

The message in (8) is that one is not better assuring a lasting marriage through intimate premarital knowledge (and up close "studying", as one does with one's opponent or prey) of one's partner. The *-mann* in (8) thus functions more indefinitely with respect to maleness or femaleness, and thereby criterion (c) is (perhaps) met.

For criterion (d), we should find a semantic shift from *-mann* to the first element in such constructions, and the meaning of the whole construction should not be directly paraphrasable. This proves difficult for Modern German *Sportsmann*, which can be paraphrased as 'one who is sporting', and it is a similar case for the other examples given in the *Duden* (1995): *Forstmann* 'one who works in forestry'; *Weltrekordmann* 'one who holds a world record'. The basic meaning can still be borne by the second constituent in these constructions, though it is possible to also construe other lexemes with just single roots, such as 'athlete', 'ranger', and 'champion' respectively.

Criterion (e) states that the affixoid is in competition with or complementary distribution with a system of affixes. The *Duden* (1995: 508-509) considers *-mann* to be in competition with the suffixes *-ler* (*Sportler* and *Weltrekordler*) and agentive *-er* (*Förster*), to follow through with the previous examples. Given the 8 examples of *-frau* in *Sportsfrau* in a google search, it may well be that this form is providing further competition. A variety of other personal suffixes exist in German, such as *-iker*, *-ist*, and *-ator*, along with those derived from status labels (*-papst* 'pope', *-baron*), animals (*-ratte* 'rat', *-hai* 'shark'), personal names (*-fritze*, *-huber*), as well as others (see also Leuschner and Wante 2009).

Having measured up the largely recognized Modern German personal suffixoid *-mann* in accordance with the five proposed tests for affixoids, a number of questions arise, including whether it is necessary for all of the criteria to be fully met in order for a suffixoid to be recognized as such, how to interpret the diachronic perspective (b, c, d), and how to interpret this phenomena in conjunction with spoken language (a, d). We turn now to these and other questions.

## 5. DISCUSSION

The application of Stevens' (2010) diagnostics, themselves largely the product of state-of-the-art consensus, yields an interesting outcome when applied to Modern German *-mann*. Some criteria seem to apply well, others are more questionable with this particular recognized suffixoid. It may well be that the class of suffixoid, similar to other linguistic classes, is not monolithic and homogeneous, but is instead made up of more central and more peripheral members. Is it in fact necessary to have all five features (a-e) be in play for a suffixoid to be recognized as such? Stevens (2010) does not explicitly make this claim, but Elsen (2009: 319) notes it is important for a constellation of factors to be on hand for affixoidization, namely productivity, semantic change, and the unbound coexistence of a cognate form. This covers criteria (a-d), leaving out (e), systemic interaction with a class of affixes. I generally find (e) useful, just as in the case of *-mann* where it seemingly helps as an identifier, but as with all or most linguistic classes, it is wise to best consider such diagnostic criteria as possibly varying in applicable degree.

Criteria (a) regarding productivity, as point (e) above, appears at first glance to be relatively unproblematic. Productivity in the modern era is established. Leuschner and Wante (2009: 66) even note that at least three constructions have been what have counted in past scholarship to meet a productivity threshold. There are in fact, however, more *-mann* constructions that have arisen in modern times (and previously), so the potential minimum question of three is moot in this case (recall the over 200 types noted above in the modern language).

Although it is stated above that criterion (b) is fulfilled, i.e., a free "parent" morph coexists from which the suffixoid is derived, there is room for discussion here. The reconstructed and historical record often show a path leading from full unbound word, to compound constituent, and then to affixoid, but I would suggest we do not wish to imply there is a strong claim that that pathway exists in isolation, i.e., out of reach of any other linguistic influences, analogy, and so-called semantic contamination. How can we assume that the development of the indefinite pronoun *man* did not also exert a semantic influence (part 4 above) and contribute to indefiniteness (in terms of gender or otherwise) and semantic generalization linked to *-mann* constructions?

A related question centers on diachronic semantics and points (b) and (c). Various *-mann* constructions have been lexicalized as compounds or suffix-like formations for at least 1,500 years, if not more than two millenia (Kluge 1999, Lightfoot 2006), and they continue to be formed anew. *Man* as an indefinite pronoun has been

recorded since at least the OHG era (*Kluge* 1999, etc.). How does one decide at which point in time a given “parent” form’s semantics ought to be scrutinized in relation to the suffixoid? It is sometimes difficult to judge if certain semantics have been shed over time. This point is particularly salient for *-mann* formations due to the modern perception of sexism and exclusion in language. An etymological approach might claim that the meaning associated with the reconstructed Germanic or OHG *\*mann* ‘man, person’ is the reflexive meaning to refer to when considering *Duden’s* (1995) modern example of *Forstmann* ‘ranger, woodsman’ (as opposed to ‘male adult’), especially since the construction goes back to OHG *forstmann* ‘woodsman’ (Henzen 1947). Indeed, it was back in that early Medieval period in which Henzen explains that the compound constituent was transitioning into something more suffixal (1947: 194-195).<sup>4</sup>

Does this mean the process of suffixoidization has been ongoing for over 1,300 years in light of the *Duden* claim that *-mann* in this construction is (still) a suffixoid? This seems less than likely for at least two reasons, which could mean speakers’ perceptions of “parent” morphs may be more salient than the actual etymological record when considering point (b).

The reasons pointing away from 1,300 years of suffixoidization include a lack of data showing similar such lengthy developments in other suffixoids, and the speaker-based nature of affixoidization. Two previous studies, examining the historical suffixoidal nature of German *-heit* ‘-hood’ as in *Kindheit* ‘childhood’ (Lightfoot 2005) and the development of the constituent *-mann* in the formation *gomman* ‘husband, man’ (Lightfoot 2006), were carried out using data from exhaustive searches in the Titus database, which contains all of the main OHG works and some of the minor ones. Evidence points toward suffixoidization having taken place for the *-heit* form within the OHG period, possibly also somewhat prior to it (cf. similar derivational forms such as *-schaft* ‘-ship’, *-tum* ‘-dom’, etc.; Erben 2000, Henzen 1947, Nübeling et al 2008). The *gomman* study (Lightfoot 2006) is mostly limited to one OHG work, Tatian’s translation of the Gospels, and it concluded somewhat more authoritatively that development from a transparent compound constituent *-man* took just one period, namely OHG, to undergo demorphologization with a host which had a similar if not the same meaning. This is not suffixoidization, but it does show the same rough amount of time that this particular form took to complete a substantial morphological change.

The second plausible reason for the relatively quick progression of most personal suffixoids is their basis in colloquial language, which tends to undergo fairly rapid change in relation to written language. *Duden* (1995: 508-509) and Stevens (2010: 73) point out this colloquial character. High productivity in spoken language tends to accelerate change (cf. the well-known rate of change associated with youth language, for example).

---

<sup>4</sup> Unfortunately, Henzen does not provide contextualized examples from the record.

This brings us back to the claim by Henzen (1947: 195) of OHG *Forstmann* being in transition already up to 1,300 years ago. He was presumably referring to there having been a generalization of the semantics of *-mann*. A definite man was not meant, and the form appeared among a list of other similar constructions. Henzen's discussion frequently mentions *second compositional constituents*, which are noted above as the second stage in the progression of root word > compound constituent > affixoid > affix. Thus I would surmise Henzen is identifying the second stage as opposed to the third.

This then begs the question of what sets apart a *second compositional constituent* from a suffixoid. Stevens (2010) does not raise this issue; his discussion emphasizes the attributes of roots, affixoids, and derivational affixes. Does this question need to be raised, or could it be the case that Henzen's concept has some sort of conceptual overlap with the category affixoid? Henzen notes at least the same qualities as (a), (c), and (d) for the *second compositional constituent* (1947: 194-195), but his proposed category seems to crucially lack the coexistence of an unbound, formally similar cognate.

This paper's assessment of (d), the shift in semantic weight from the second element to the first, along with the inability to paraphrase the whole construction by means of its two related root meanings, with respect to the *-mann* examples is problematic, as pointed out above in part 4. By and large, for *Sportsmann*, it almost always is a male referred to who is sporting, athletic, or has the positive qualities of a sporting man. Likewise a *Weltrekordmann* refers to a man who is a champion or has a world record. On the other hand, (d) seems perfectly in order in applying to other sorts of suffixoids such as *-werk*. In *Laubwerk* 'foliage' (*Laub* 'leaves'), for example, foliage is not "a work of leaves", but a complex whole of greenery or a collective entirety of greenery. In such cases the paraphrasing using the two original roots' meanings does not work, and indeed the semantic weight ends up with the first element.

Finally, the challenges with *-mann* and the prescribed semantic relations, changes, and paraphrasing lead one to consider whether this is a good example of an affixoid at all. It may alternatively be that some affixoids are better members of this category, and some are more peripheral members. This seems more likely than the possibility that Stevens' (2010) tests are simply off the mark. It has already been pointed out in the literature that personal suffixoids have a particular colloquial connection, and we can guess that this influences their usage and development. The form *-mann* itself actually does not appear in Leuschner and Wante's (2009) study of personal suffixoids, nor does it appear in Elsen's (2009) work, so this may lend weight to the argument that *-mann* is not an ideal candidate for recognition as a suffixoid, despite its presentation as such in the *Duden* (1995).

## 6. CONCLUSION

This paper examined the tests for affixoids proposed by Stevens (2010) and utilized them to test the German personal suffixoid *-mann*. The test points themselves are deemed to be workable. The results here lead to a number of questions and thoughts about the tests themselves, including whether all must be satisfied to count as a litmus test for membership. This work also raises issues such as whether differ-



ent subclasses of affixoids need to be considered, and questions whether Modern German *-mann* actually belongs categorized in the *Duden* (1995: 508) as a suffixoid.

Previous literature on the topic has not emphasized that the colloquial origin (Stevens 2010: 73) of suffixoids is not to be forgotten—this can make looking for them in the written record especially problematic, and is an important methodological consideration, reminding the researcher to be extremely cautious when dealing with historical data. Personal suffixoids in particular (*Duden* 1995) occur perhaps even more so in colloquial domains, so written dialogue and personal electronic communication such as blogs are likely fruitful places for their further investigation, given the oral influence presumed in those domains.

### PRIMARY SOURCES

- Google. 7 September 2011. <http://google.de>.  
Mypashion. 7 September 2011. <http://www.mypashion.com>  
*Der Spiegel*. 7 September 2011. <http://www.spiegel.de>.  
Titus database. 29 August 2011. <http://titus.uni-frankfurt.de>.

### REFERENCES

- BOOIJ, Geert (2009) “Compounding and construction morphology.” In: R. Lieber/P. Stekauer (eds), 201-216.  
ELSEN, Hilke (2009) “Affixoids: only what is named can also be understood.” *Deutsche Sprache* 37/4, 316-333.  
ERBEN, Johannes (2000) *Einführung in die deutsche Wortbildungslehre*. 4<sup>th</sup> ed. Berlin: Erich Schmidt Verlag.  
*DUDEN* (1995) *Grammatik der deutschen Gegenwartssprache*. P. Eisenberg, et al (eds). Mannheim: Dudenverlag.  
HENZEN, Walter (1947) *Deutsche Wortbildung*. Halle: Max Niemeyer Verlag.  
KASTOVSKY, Dieter (2009) “Diachronic perspectives.” In: R. Lieber/P. Štekauer (eds), 323-340.  
*KLUGE* (1999) *Etymologisches Wörterbuch der deutschen Sprache*. E. Seebold, et al (eds). Berlin: Walter de Gruyter.  
LEUSCHNER, Torsten/Tanja MORTELMANS/Sarah DE GROOT (eds) (2010 [2005]) *Grammatikalisierung im Deutschen*. Berlin: Walter de Gruyter.  
LEUSCHNER, Torsten/Eline WANTE (2009) “Personale Suffixoide im Deutschen und Niederländischen.” *Germanistische Mitteilungen* 70, 59-73.  
LIEBER, Rochelle/Pavol STEKAUER (eds) (2009) *The Oxford Handbook of Compounding*. Oxford: Oxford University Press.  
LIGHTFOOT, Douglas (2000) *On the Grammaticalization of a Nominalizing Suffix in German: -tum*. Dissertation. University of California, Los Angeles.  
LIGHTFOOT, Douglas (2005) “Can the lexicalization/grammaticalization distinction be recociled?” *Studies in Language* 29/3, 583-615.

- LIGHTFOOT, Douglas (2006) "Aspects of the rise and fall of the compound *gomman* in *Tatian*." *Leuvense Bijdragen* 95, 71-86.
- LIGHTFOOT, Douglas (2010 [2005]) "Gestalt derivation in German." In: T. Leuschner/T. Mortelmans/S. De Groot (eds), 59-70.
- MATER, Erich (1967) *Rückläufiges Wörterbuch der deutschen Gegenwartssprache*. Leipzig: VEB Verlag Enzyklopädie.
- MATTHEWS, P.H. (1991) *Morphology*. 2<sup>nd</sup> ed. Cambridge: Cambridge University Press.
- NÜBLING, Damaris, et al (2008) *Historische Sprachwissenschaft des Deutschen. Eine Einführung in die Prinzipien des Sprachwandels*. 2<sup>nd</sup> ed. Tübingen: Gunter Narr Verlag.
- SCHMIDT, Günther Dietrich (1987) "Das Affixoid: zur Notwendigkeit und Brauchbarkeit eines beliebten Zwischenbegriffs der Wortbildung." In: R. Wimmer/G. Zifonun (eds), *Deutsche Lehnwortbildung*, 53-101. Tübingen: Gunter Narr Verlag.
- STEVENS, Christopher (2000) "The derivational suffixes and suffixoids of Old Saxon. A panchronic approach to a linguistic category." *American Journal of Germanic Linguistics and Literatures* 12, 53-79.
- STEVENS, Christopher (2010 [2005]) "Revisiting the affixoid debate: on the grammaticalization of the word." In: T. Leuschner/T. Mortelmans/S. De Groot (eds), 71-83.

#### Abstract

#### TESTING THE SUFFIXOIDIZATION OF GERMAN *-MANN* 'man'

Throughout its history in Old High German (OHG), Middle High German, Early New High German, and Modern German (ModG), the word *Mann* 'man' has exhibited a high degree of productivity and linguistic flexibility. It has ranged morphosyntactically from a full noun (OHG *man* 'man'), compounded noun (OHG *werolt.man* 'human, mortal'), affixoid (ModG *Sports.mann* 'athlete'), indefinite pronoun (ModG *man* 'one'), to something approaching zero (undergoing demorphologization in the OHG compound *gom.man* 'husband, man'). The *affixoid* notion and its usefulness in morphology have been controversial in the literature (e.g., Schmidt 1987). Relatively recently, Stevens (2005) proposed criteria for the category *affixoid*, and a number of standard German grammatical and etymological references (e.g., *Duden* 1995) have been utilizing this term. This study involves examining modern usage of the form *-mann* as found in the online *Spiegel* newsmagazine's database, as well as the collection of historical data primarily on the forms of German *-mann* from the robust *Titus* database in Frankfurt. Stevens' (2005) criteria are evaluated and used to measure the validity of German *-mann*'s membership in the category of *suffixoid*.

Povzetek  
PREVERJANJE SUFIKSOIDIZACIJE NEMŠKEGA »-MANN« 'moški'

V svoji zgodovini od stare visoke nemščine (SVN), srednje visoke nemščine, zgodnje nove visoke nemščine in moderne nemščine (MN) je beseda *Mann* 'moški' poznala različne stopnje produktivnosti in jezikovne spremenljivosti. Njen morfosintaktični status je segal od polnopomenske besede (SVN *man* 'moški'), zložene besede (SVN *werolt.man* 'človek, smrtnik'), afiksoida (MN *sports.mann* 'športnik'), nedoločnega zaimka (MN *man* 'nekdo') do skoraj ničtega statusa (napredujoča demorfologizacija v zloženki (SVN) *gom.man* 'mož, soprog'). Pojem *afiksoida* in njegova koristnost za morfologijo sta v znanstveni literaturi izzvala številne polemike (glej npr. Schmidt 1987). Šele nedolgo nazaj je Stevens (2005) predlagal kriterije za določanje kategorije afiksoida in številne priznane nemški slovnice in etimološki slovarji so termin pričeli uporabljati. Naša razprava bo raziskovala moderno rabo oblike *-mann*, kakršno najdemo v spletnem arhivu tednika *Spiegel* in v historičnih podatkih o oblikah nemškega *-mann* iz bogate podatkovne zbirke *Titus* iz Frankfurta. Ovrednotili smo Stevensove kriterije (Stevens 2005) in jih uporabili za določanje upravičenosti umestitve nemškega *-mann* v kategorijo *sufiksoidov*.



## DERIVATION IN A WORD-BASED MORPHOLOGY: ON THE ORIGIN OF OLD ENGLISH VERBS OF PERCEPTION, COGNITION AND EMOTION

### INTRODUCTION

This paper deals with some methodological and theoretical issues raised by diachronic lexical morphology (Pena 2008; Trips 2009). More exactly, in this paper I am going to describe the process of lexicalization in the grammaticalization of a selection of Old English causative verbs through the analysis of some of the different strategies for the expression of causation in three different lexical domains: PHYSICAL PERCEPTION, COGNITION and EMOTION. I will argue here that speakers can express relative measures of causation and that these grammatical variations are clearly reflected in the verb's lexical strata (Giegerich 1999: 87-88). Furthermore, I will describe some of the different ways synchronic patterns of morphological variation are to be related to diachronic processes of lexical formation, in order to show the relationship between the form and the meaning of this set of words.

My first aim will be to identify some of the different ways the grammatical parameter of causation finds a systematic correspondence on the syntax of the Old English verb, acting as a filter that accepts certain morphological patterns while blocking others. Furthermore, I will try to show that the strategies used in OE for the expression of causation can vary from one lexical domain to another. This implies that whereas all the predicates from the same lexical domain follow the same set of derivational processes, very significant differences can be detected when predicates from different domains are analyzed. In doing so, I want to show the existence of different grammars of causation, which broadly speaking correspond to the different degrees of causativity each predicate can attain.

### 2. THE 'CAUSATION IS FORCED MOTION' METAPHOR

Causation is a basic human concept. However, this does not imply that causation is an undecomposable primitive. Lakoff and Johnson (1999) state that human beings interpret the concept of CAUSE through conceptual metaphors, and many of these metaphors are related to MOTION. If, for example, object A moves towards object B setting it in motion, we assume that, without object A's movement, object B would have remained in its original position. The later extension of this understanding of causation to other domains, physical or abstract, implies a change towards the general identification between CAUSATION and FORCED MOTION. According to Lakoff and

---

\* *Author's address:* Universidad de Castilla-La Mancha, Facultad de Letras, Filología Moderna, Avda. de Camilo José Cela s/n, 3071 Ciudad Real, España. Email: JavierEnrique.Diaz@uclm.es

Johnson (2003: 69-70), causation is best understood as a cluster of components that form a gestalt:

We can see this most clearly in infants. Piaget has hypothesized that infants first learn about causation by realizing that they can directly manipulate objects around them—pull off their blankets, throw their bottles, drop toys. There is, in fact, a stage in which infants seem to »practice« these manipulations, e.g., they repeatedly drop their spoons. Such direct manipulations, even on the part of infants, involve certain shared features that characterize the notion of direct causation that is so integral a part of our constant everyday functioning in our environment—as when we flip light switches, button our shirts, open doors, etc. Though each of these actions is different, the over-whelming proportion of them share features of what we may call a »prototypical« or »paradigmatic« case of direct causation.

The underlying metaphor CAUSATION IS FORCED MOTION and its reflections in the derivational morphology of Proto-Germanic have been described in Díaz Vera (2000). In that paper, I argued that two of the three basic derivational patterns used by speakers of Primitive Old English (and other ancient Germanic languages) in order to derive causative verbs from non-causative ones are based on the combination of the corresponding lexical root (normally a noun) plus the formative \*/jα/. Furthermore, it is argued that this formative was originally used in order to derive agents from nouns (Lowe 1972: 215-215), as in the following examples:

PrGmc \*herd- ‘herd’ + \*jα- > PrGmc \*herdjα- > OE *herde* ‘a keeper of a herd; someone who guides a herd’

PrGmc \*do:m- ‘a law, statute’ + \*jα- > PrGmc \*do:mjα > OE *dēma* ‘a judge, arbiter, ruler; someone who administers law’

According to Pokorny (1959-69: 1.293), the ultimate origins of this Germanic suffix are to be found in the Indo-European verb \*/ei-/, with the meaning ‘to cause something to move’, illustrated by the Greek verb εἶμι ‘to put in motion’, which came to express ‘motion in the future’ in Attic and in Ionian (Liddell and Scott 1968: 489). In the examples above, the causative character of this formative is clear in the first example (OE *herde* ‘someone who guides a herd’) and, more figuratively, in the second one (OE *dēma* ‘someone who administers law’).

This derivational morpheme is also found with verbs expressing motion, as in the case of the Primitive OE weak verb \*du:fjα-n (OE *dýfan*) ‘to immerse, to cause something to move by diving’, derived from the strong verb \*du:foxn (OE *dūfan*) ‘to dive’. This basic construction reflects a situation where a causer (i.e. the subject/agent) has a direct influence on the causee (i.e. the object), which will change its position due to the causer’s action. Furthermore, the rapid expansion of this derivational morpheme to other verbs from different lexical domains, both physical and abstract, and to nouns referring to different bodily organs (both external and internal) implied the develop-

ment of a brand-new morphological category of verbs, normally known as ‘weak verbs’, and the progressive demise of the ablaut system inherited from Indo-European.

### 3. THE FORMATIVE \*JA- COMBINED WITH NOUNS FOR EXTERNAL BODILY ORGANS

Most Old English verbs of physical perception are denominal, that is, they are derived from the noun of the corresponding bodily organ, to which the formative \*/jα/ ‘to (cause to) move’ is added. This is especially true for the following set of verbs of HEARING, SMELLING, TOUCHING and TASTING:

PrGmc \*αuz- ‘ear’ + \*jα- > PrGmc \*hαuzjan- > OE *hieran* ‘to hear’

PrGmc \*nαs- ‘nose’ + \*jα- > PrGmc \*nαsjan > OE \**næsian* > ME *nese* ‘to smell’

PrGmc \*fo:l- ‘hand, palm’ + \*jα- > PrGmc \*fo:ljan- > OE *fēlan* ‘to touch’

PrGmc \*beuk- ‘belly, body, trunk’ + \*jα- > PrGmc \*beurjan > OE *byrgan* ‘to taste’

Only in the case of Old English verbs expressing SEEING, I have not been able to find illustrations of this derivational pattern. However, the OE verb *scēawian* ‘to look’, which derives from PrGmc \*skawo:jan (Orel 2003: 337), a certain degree of phonetic similitude with the noun for the corresponding sensory organ (PrGmc \*αug- > OE *ēage* ‘eye’) can be detected.

In this second folk model of causation, we have a situation where a causer (i.e. the entity from which the physical sensation emanates) has a direct influence on the causee (i.e. the corresponding external organ in the perceiver’s body), which will change its form or texture due to the causer’s action. This interpretation is fully coherent with the ancient seal-in-wax metaphor (Carruthers 1990: 21-30), according to which the act of perceiving produces a change in the perceiver; this change can be material (as in the case of TOUCHING) or spiritual (as in the other four cases). As Carruthers (1990: 70-71) puts it:

What the eye takes from the object via air and light is its ‘formal’ aspect but not its ‘material’ one, and this aspect causes the eye to change in the sensation we call ‘sight’. The change in the eye occurs in the same manner in which phantasms are recorded in memory.

### 4. THE FORMATIVE \*JA- COMBINED WITH NOUNS FOR INTERNAL BODILY ORGANS

The development of causative verbs of FEELING follows the morphological pattern described above for verbs of PHYSICAL PERCEPTION. However, verbs of FEELING are derived from nouns referring to the internal organ that is considered the see of each feeling in Anglo-Saxon folk beliefs. This is the case of OE *heorte* ‘heart’, which is considered the see of courage and fierceness:

PrGmc \*hert- ‘heart’ + \*-jα- > PrGmc \*hertjαn- > OE *hiertan* ‘to cause to have courage, to encourage’

This construction triggers a third folk model of causation as possession, where a causer (i.e. the entity from which the emotion emanates) takes hold of the causee (i.e. the corresponding internal organ in the experiencer’s body), which will change its ownership due to the causer’s action. In turn, possession can be conceptualized in terms of location, so that something (an emotion) is in a certain place (a bodily organ). Similarly, we have in Old English the following verb:

PrGmc \*mun- ‘mind’ + \*-jα- > PrGmc \*munjαn- > OE *gemynan* ‘to cause to remember, to reminisce’

PrGmc \*man- ‘mind’ + \*-jα- > PrGmc \*manjαn- > OE *gemanian* ‘to cause to remember, to reminisce’

Differently to *thought* (Ide \*tαng- > OE *ðyncan*), the *mind* (Ide \* men- > OE *munan*) was related to a concrete bodily organ (either the heart or the brain). Our evidence from Old English and other Germanic languages clearly indicates that whereas *memory* was considered a function of a bodily organ, *thinking* was metaphorically mapped on the more abstract area of FEELING.

As for verbs of EMOTIONS, OE *lufian* ‘to love’ illustrates the construction described in this paper:

PrGmc \*liβr- ‘liver’ + \*o-jα- > PrGmc \*luβojαn- > OE *lufian* ‘to cause to feel love, affection’

Differently to the other Old English verbs described here, OE *lufian* is a Class II weak verb, which is composed of deverbal and denominal verbs with what is normally known as an intensive or iterative sense. From a chronological point of view, these verbs are relatively newer than Class I weak verbs, as reflected by the fact that their root vowels did not undergo the effects of /i/-Umlaut (Görlach 1997: 84). It has been suggested that causativity and intensive/iterative are two aspects of one original morpheme as, for example, in Indo-European (where both are expressed by the same suffix \*/ei-/; Lehmann 1993: 168), in Old Chinese (Schuessler 2007: 19) and in Amerindian languages (Broadwell 1997). Comrie (1985: 330) hypothesizes that this fact may be related to the “intensity (force) that is often implied in forcing someone to do something.”

Causative verbs expressing more abstract emotions, such as *sadness*, *grief*, *repent* or *pleasure*, do not have a clear identification with a concrete bodily organ. Consequently, rather than from a noun referring to a bodily organ, the corresponding verbs are derived from the name of each emotion, as can be seen from the following examples:

PrGmc \*αgl- ‘sadness’ + \*-jα- > PrGmc \*αgljαn- > OE *eglan* ‘to cause to feel sadness, affliction’



PrGmc \*lust- ‘pleasure’ + \*-jα- > PrGmc \*lustjαn- > OE *lystan* ‘to cause to feel pleasure’

PrGmc \*skām- ‘shame’ + \*-jα- > PrGmc \*skāmjαn- > OE *scēamian* ‘to cause to feel shame’

PrGmc \*hrew- ‘regret’ + \*-jα- > PrGmc \*hrewjαn- > OE *hreowian* ‘to cause to feel regret, repentance’

## 5. CONCLUSIONS

In this brief analysis I have tried to show the following:

1. The basic causative morpheme \*-jα-, with the meaning ‘to cause to move’, was used to derive a whole new verbal class in Old English.
2. However, when applied to verbs of PERCEPTION, FEELING and EMOTION, the folk model of cause that underlies these constructions seems to see the world in terms of the naturalness of events, rather than in terms of causation proper.
3. In the case of PERCEPTION verbs, the causee corresponds to a sensory organ (nose, ears, stomach or skin), which undergoes a change (physical or abstract) in its form as a consequence of direct contact with the cause, which corresponds to the stimulus.
4. Whereas verbs of cognition (such as OE *gemynan*) follow a similar schema, the conceptualization represented by some Old English verbs of emotion combine causativity and iteration meanings. The emotion emanates directly from the bodily organ (as in the case of OE *hiertan* and *lufian*), changing its substance and causing the person to react.

## References

- BROADWELL, George A. (1997) “When causative means ‘intensive’.” Presented at the Society for the Study of the Indigenous Languages of the Americas summer meeting, Ithaca, NY July 15, 1997.
- CARRUTHERS, MARY (1990) *The Book of Memory: A Study of Memory in Medieval Culture*. Cambridge: Cambridge University Press.
- COETSEM, Frans van/Herbert L. KUFNER (eds) (1972) *Towards a Grammar of Proto-Germanic*. Berlin: Mouton de Gruyter.
- COMRIE, Bernard (1985) “Causative verb formation and other verb-deriving morphology.” In T. Shopen (ed.), 303-345.
- DÍAZ VERA, Javier (2000) “The development of causation in Old English and its interaction with lexical and syntactic diachronic processes.” *Cuaderno de Investigaciones Filológicas* 26, 17-38.
- GIEGERICH, Heinz (1999) *Lexical Strata in English: Morphological Causes, Phonological Effects*. Cambridge: Cambridge University Press.
- GOÖRLACH, Manfred (1997) *The Linguistic History of English: An Introduction*. London: Palgrave.
- LAKOFF, George/Mark JOHNSON (1999) *Philosophy in the Flesh*. Chicago: University of Chicago Press.

- LAKOFF, George/Mark JOHNSON (2003) *Metaphors We Live By*. Chicago: University of Chicago Press.
- LEHMANN, Winfred, P (1993) *Theoretical Bases of Indo-European Linguistics*. London: Routledge
- LIDDELL, Henry George/Robert SCOTT (1968) *Greek-English lexicon*. Oxford: Clarendon.
- LOWE, Pardee Jr (1972) "Germanic word formation." In: F. van Coetsem/H.L. Kufner (eds), 211-238.
- OREL, Vladimir (2003) *A Handbook of Germanic Etymology*. Leiden/Boston: Brill.
- PENA, Jesús (2008) "El cambio morfológico en el interior de las series de derivación." *Revista de Investigación Lingüística* 11, 233-248.
- POKORNY, Julius (1959-69) *Indogermanische Etymologisches Wörterbuch*. Bern/München: Francke Verlag.
- SCHUESSLER, Alex (2007) *ABC Etymological Dictionary of Old Chinese*. Honolulu: University of Hawai'i Press.
- SHOPEN, Timothy (ed) (1985) *Language typology and syntactic description*. Cambridge: Cambridge University Press.
- TRIPS, Carola (2009) *Lexical Semantics and Diachronic Morphology - The Development of -hood, -dom and -ship in the History of English* (Linguistische Arbeiten 527). Tübingen: Max Niemeyer Verlag.

#### Abstract

### DERIVATION IN A WORD-BASED MORPHOLOGY: ON THE ORIGIN OF OLD ENGLISH VERBS OF *PERCEPTION*, *COGNITION* AND *EMOTION*

In this paper I describe the process of grammaticalization of Old English causative verbs through the analysis of some of the different strategies for the expression of causation in three different lexical domains: PHYSICAL PERCEPTION, COGNITION and EMOTION. The etymological analysis of these verbs indicates that, in the case of the three lexical domains under scrutiny, causation was lexicalized and conceptualized as forced movement or change of an external (as in the case of verbs of FEELING; e.g. OE *hieran* 'to hear') or internal (i.e. verbs of COGNITION and some EMOTION verbs; e.g. OE *gemynan* 'to remember' and OE *hiertan* 'to encourage') part of the body.

#### Povzetek

### IZPELJEVANJE V BESEDNOUTEMELJENI MORFOLOGIJI: O IZVORU STAROANGLEŠKIH GLAGOLOV *ZAZNAVE*, *SPOZNAVANJA* IN *ČUSTEV*

V pričujočem članku proces gramatikalizacije staroangleških kavzativnih glagolov opisujem s pomočjo analize nekaterih strategij izražanja kavzalnosti na treh različnih leksikalnih ravneh: FIZIČNA ZAZNAVA, SPOZNAVANJE in ČUSTVA. Etimološka analiza teh glagolov pokaže, da je bila v primeru analiziranih leksikalnih področij kavzalnost leksikalizirana in konceptualizirana kot vsiljeno gibanje oziroma sprememba zunanjega (kot na primer pri glagolih *OBČUTIJ*, npr. staroangleški *hieran* 'slišati') ali notranjega (na primer glagoli *SPOZNAVANJA* in nekateri glagoli *ČUSTEV*, npr. staroangleška *gemynan* 'spominjati se' in *hiertan* 'spodbujati') telesnega dela.

**THE ARCHAIC WORD-FORMATIONAL RELATIONSHIP  
BETWEEN THE PROTO-SLAVIC NOUN \*POL'E  
AND THE HITTITE ADJECTIVE PALHI-**

1. In his 1982 article, Calvert Watkins demonstrated that the PIE acrostatic and proterokinetic declension patterns of *u*-stems originally expressed a functional relationship, in which nouns were declined following the acrostatic type and the word-formationally identical adjectives following the proterokinetic type,<sup>1</sup> and that the traces of such a relationship can still be recognized, for example, in Vedic Sanskrit, in which the noun *vásu-* (n.) 'good' is acrostatically inflected (cf. gen. sg. *vásvah*), whereas the homophone adjective *vásu-* 'good' is proterokinetically inflected (cf. gen. sg. *vásoh*).<sup>2</sup> Based on the archaic Hit. nom-acc. pl. n. *a-aš-šu-u* 'good' < \**-u-H*<sub>2</sub> (alongside standard *āššu<sub>u</sub>* 'idem'), which inflectionally corresponds to Ved. *vásū* (nom-acc. pl. n.) 'good' and Av. *vohū* 'idem', Watkins therefore postulated that the word-formational relationship between the Hit. adjective *āššu-/āšša<sub>u</sub>* 'good, dear, favored' and the noun *āššu-* (n.) 'good' is not as is generally explained, as though the noun is substantivized from the adjective, but that the noun *āššu-* (n.) 'good' was inflected in the protolanguage following the acrostatic accent pattern and has therefore also preserved the archaic case ending \**-H*<sub>2</sub> in the nom-acc. pl. n. The proterokinetic quality of *-u/au-* in the adjective *āššu-/āšša<sub>u</sub>* and the acrostatic quality of the noun *āššu-* (n.) 'good' with the archaic form *a-aš-šu-u* (nom-acc. pl. n.) thus presumably still expresses the protolanguage state of affairs.

2. Watkins' explanation of the original functional relationship between the acrostatic and proterokinetic accent-ablaut pattern can be typologically compared with the situation in more recent thematic relationships of the type φόρος (m.) : φορός (adj.) and may therefore merely be a transfer of the older protolanguage pattern:

acrostatic paradigm → nominal declension

proterokinetic paradigm → adjectival declension

Indirect indicators of this protolanguage pattern may include relationships of the type Skt. *ápas-* (n.) 'work' : *apás-* (adj.) 'active' and also the more recent type Skt. *uṣr-á-* (adj.) 'morning, reddish' : *uṣár-* (f.) 'sunrise, morning', *tamasá-* (adj.) 'dark' : *támas-*

---

\* *Author's address:* Inštitut za slovenski jezik Frana Ramovša, Novi trg 4, 1000 Ljubljana, Slovenia.  
Email: metka.furlan@zrc-sazu.si

<sup>1</sup> A similar interpretation of the functional relationship between the acrostatic and proterokinetic declension of *i*-stems and *u*-stems is also found in Benveniste (1935: 52).

<sup>2</sup> Following Melchert (1994a: 301) and Pinault (2003).

(n.) ‘darkness’. The last example of adjective formation includes not only thematization of a consonant noun but (also) shift of the accent onto it. Even substantivization with an accentual shift of the type Skt. *kṛṣṇá-* (adj.) ‘black, dark’ → *kṛṣṇa-* (n.) ‘black antelope’, Gr. λευκός (adj.) ‘bright shining’ → λεῦκος (m.) ‘grey mullet’ could be an expression of this protolanguage word-formational relationship.

The original functional relationship between the acrostatic and proterokinetic declension patterns must have started to break down very early in the protolanguage, before the Anatolian branch separated from the protolanguage system. In Hittite, *i*-stem and *u*-stem nouns are *predominantly* inflected following a declension pattern without ablaut; that is, acrostatically (e.g., *ḫalki-š*, *ḫalki-n*, *ḫalkiḫ-aš* etc.; *\*ḫaššu-š*, *\*ḫaššu-n*, *\*ḫaššuḫ-aš* etc.), and *i*-stem and *u*-stem adjectives are *predominantly* inflected following a declension pattern with ablaut; that is, proterokinetically (e.g., *šalli-š*, *šalli-n*, *šallaš* (< *\*šallaḫ-aš*) etc.; *āššu-š*, *āššu-n*, *āššuḫ-aš* etc.). Hittite preserved this protolanguage feature as a tendency with a number of exceptions among both nouns and adjectives; for example, *ḡeši-* (c.) ‘pasture’: *ḡešaeš* (nom. pl.), *ḡešauš* (acc. pl.); NINDA*ḫarši-* (c.) ‘type of bread’: NINDA*ḫar-ša-i* (dat.-loc. sg.), NINDA*ḫar-ša-eš* (nom. pl.), NINDA*ḫar-ša-uš* (acc. pl.); DUG*palḫi-* (n.) ‘type of (broad) vessel’: DUG*pal-ḫa-aš* (dat.-loc. pl.), DUG*pal-ḫa-e-a*<sup>HIA</sup> (nom.-acc. pl.); *šēli-* (c.) ‘pile of grain, granary (?)’: *še-e-la-aš* (gen. sg.) etc.; *nakki-* (adj.) ‘important; difficult’: *nakkiḫ-az* (abl.), *nakki-uš* (acc. pl. c.) etc.

There still exists an inherited functional distinction in Hittite between the homophone noun *āššu-* (n.) and the adjective *āššu-/āššuḫ-*, but not between the homophone pair DUG*palḫi-* (n.) : *palḫi-* (adj.), from which it can be concluded that the second pair is of more recent origin because the noun could have arisen via the “zero substantivization” type Skt. *pṛthivī-* (f.) ‘earth’ ← *pṛthivī-* (adj. f.). It was probably the zero substantivization pattern that triggered the breakdown of the original relationship between the protolanguage acrostatic and proterokinetic declension types, with the result that the originally acrostatically declined nouns began switch over to the proterokinetic type, such as *\*deru-* (n.) ‘wood, tree’, in which the original acrostatic pattern, alongside Skt. and Gr. material, is also confirmed by PSI. *\*deṛvo* (n.), which arose with the thematization of the oblique case stem *\*deru-*:<sup>3</sup>

nom.-acc. sg. *\*dóru* = Skt. *dāru*, Gr. *δόρυ*, Hit. *tāru-*  
 gen. sg. *\*déru-s* (→ *\*déru-o-m* > PSI. *\*deṛvo*) → *\*dréu-s* = Skt. *dróḥ* (→ *\*dréu-o-m* =  
 Goth. *triu*)

<sup>3</sup> With the same thematization from PIE *\*seru-/soru-* (n.) ‘quarry’ (cf. Hit. *šāru-* (n.) ‘quarry’, OIr. *serb* (f.) ‘theft’ < *\*serueH<sub>2</sub>*), it would be possible to explain membership in this word family of Lat. *seruus* (m.) ‘slave’ (EIEC: 77), in which the change of grammatical gender was probably influenced by the narrowing of the meaning to pertain only to human quarry: ‘quarry’ → *\*‘quarry = people’* → ‘slave’.

The transition of the nominal acrostatic pattern into the proterokinetic one in \*g'enu- (n.) 'knee' alongside Goth. *kniu* (n.) < \*g'n-éu-o-m, as is known, with the archaic instrumental *ga-nu-ut* < \*g'n-éu-d is also confirmed by Hittite:

nom.-acc. sg. \*g'ónu = Skt. *jānu*, Av. *zānu*, Gr. *γόνυ*  
 gen. sg. \*g'énu-s (→ Hit. *gēnu-*, Lat. *genū*) → \*g'néu-s = Skt. *jñóh* (Hit. *ganut*) (→  
 \*g'néu-o-m = Goth. *kniu*)

This phenomenon almost completely obscured the original functionality of the acrostatic and proterokinetic declension pattern. It appears to still be preserved in the Slavic-Anatolian relationship \*pòl'e (n.) : *palhi-* (adj.).

3. After Julius Pokorny first demonstrated in his *Indogermanisches etymologisches Wörterbuch* that only the Slavic neuter noun \*pòl'e 'field' and the Hittite adjective *palhi-* 'wide' – as part of a word family that he combined under the root \*pelə-/plā- 'wide and flat; to spread out; to make flat by pressing and beating, to beat out, to pat out' (IEW: 805 ff.)<sup>4</sup> – contain an *i*-stem morpheme, Schmitt-Brandt (1967: 71) derived the Hit. adjective from the reduced grade \*p<sub>e</sub>lHi-, and the Slavic noun, understandably, from the *o*-stem \*polH<sub>i</sub>on or, better, \*pólH<sub>2</sub>i-o-m.<sup>5</sup> I myself (Furlan 1986: 96; 1994: 11) have also written about this shared Slavic-Anatolian morphological feature and have demonstrated that the PSl. noun \*pòl'e (n.) was thematized following the same pattern as PSl. \*mòr'e (n.) 'sea' from a PIE neuter *i*-stem noun; specifically, \*pòl'e (n.) from PIE \*pólH<sub>2</sub>-i- (n.), and PSl. \*mòr'e from PIE \*mór-i- (n.), cf. OIr. *muir*, Lat. *mare*.<sup>6</sup> The PSl. nouns \*pòl'e (n.) and \*mòr'e (n.) are therefore, like PSl. \*deřvo (n.), thematized nouns of the acrostatic type, only that \*deřvo preserves the oblique case stem, but \*pòl'e (n.) and \*mòr'e a direct case stem.

In the same papers, I also surmised that the PIE noun \*pólH<sub>2</sub>-i- (n.) arose through substantivization from a PIE adjective, which is preserved in Hit. *palhi-* 'wide', and that this adjective therefore derived from the same *o*-grade form \*polH<sub>2</sub>-i.<sup>7</sup> The explanation is based on the presumption that the Slavic-Anatolian relationship \*pòl'e (n.) : *palhi-* (adj.) derives from the original PIE adjective \*pólH<sub>2</sub>-i- 'wide', which was nominalized in PIE and then thematized in the PIE → PSl. transition, whereas in Hittite it was preserved in its original adjectival form as a reflex of PIE \*pólH<sub>2</sub>-i.

By naming parts of the earth's surface \*pòl'e, the Slavs therefore marked it with the seme *width/extension*. This Slavic lexeme is etymologically (in its semantic moti-

<sup>4</sup> Prior to this, Benveniste (1935: 151) connected this word family or Lat. *plānus* only with Hit. *palhi-* and demonstrated that it represents the full grade \*pél-ə<sub>2</sub>.

<sup>5</sup> This explicit connection was not observed in the literature on Hittite, nor was it observed in Derksen (2008), where only root-related Germ. *Feld* is cited alongside PSl. \*pòl'e. This has been known in Slovenian etymology since the eighteenth century in the work of Marko Pohlin.

<sup>6</sup> On the secondary character of vowel coloring in Lat., see Schrijver (1991).

<sup>7</sup> This explanation of mine was later used in *ESSJ* (III, 82) and adopted in *ESJS* (678).

vation) determined exactly the same as, for example, Skt. *pṛthivī́* (f.) ‘earth’, which is originally the nominalized feminine adjective *pṛthivī́* (adj.) ‘wide, extended’ from *pṛthú-* (adj.) ‘wide, extended’.

Although such an etymological explanation of PSl. *\*pǝl'e* is also likely because the same PIE word family also includes the semantically comparable but word-formationally different Germ. *Feld* (n.) ‘field, area’, OHG *feld* ‘field, plain’ < WGerm. *\*felPa-* (n.) < PIE *\*pélH<sub>2</sub>-to-m* or Arm. *hoł* ‘earth, land, ground’ < PIE *\*polH<sub>2</sub>-o-s* (IEW: 805 ff.; Olsen 1999: 53, 781),<sup>8</sup> today it is not possible to agree with the morphological part of the explanation, that PIE *\*pólH<sub>2</sub>-i-* (n.) → PSl. *\*pǝl'e* was substantivized from the PIE adjective *\*pólH<sub>2</sub>-i-* ‘wide’, because the assumptions about PIE substantivization cannot be satisfactorily argued. Namely, neither of the other two thematized nouns *\*mǝr'e* and *\*deřvo* have such an adjectival member in their word family. Internal Hittite evidence also leads to the more economical interpretation of the pair *\*pǝl'e* (n.) : *palhi-* (adj.) with the application of Watkins’ model, which does not permit derivation of Hit. *palhi-* (adj.) from *o*-grade *\*pólH<sub>2</sub>-i-*.

Because of the Hit. proterokinetic declension type of the adjective *palhi-* (cf. *palhi-iš* (nom. sg. c.), *pal-ha-a-e-eš* (nom. pl. c.)), and especially because of the possibility that assimilation was at work in PIE *\*-VIH<sub>2</sub>V-* > Hit. *-VIV-*,<sup>9</sup> if the adjective is derived from *\*pólH<sub>2</sub>-i-*<sup>10</sup> > *\*\*palli-*<sup>11</sup> then it is more likely that *palhi-* ‘wide’ preserves a base from the PIE zero-grade root *\*pǝH<sub>2</sub>-i-*,<sup>12</sup> which became independent from *\*pǝH<sub>2</sub>-éǝ-*<sup>13</sup> and therefore derives from the adjectival proterokinetic declension pattern. However, in the direct cases, this contained the stem *\*pólH<sub>2</sub>-i-*, taking into account that the

<sup>8</sup> At least three root-connected examples (i.e., PSl. *\*pǝl'e*, Germ. *\*felPa-*, and Arm. *hoł*) show that Pokorny’s reconstruction of the root *\*pelǝ-/\*plǝ-* = *\*pelH<sub>2</sub>-/\*pleH<sub>2</sub>-* was justified.

<sup>9</sup> The assimilation *\*-VRHV-* > Hit. *-VRRV-* took place before the transition of PAnat. syllabic sonants into the Hittite reflexes *vowel + sonant* (Melchert 1994b: 55).

<sup>10</sup> At one time Couvreur (1937: 216) excluded the possibility of *o*-grade ablaut in *palhi-* because he determined that in such a case it would be written as *\*pa-al-hi-* and not *pal-hǝ*, and so he derived the adjective from a zero-grade form; that is, *\*pǝH<sub>2</sub>-i-*.

<sup>11</sup> Also taking into consideration the possibility that the assimilation *\*-VIH<sub>2</sub>V-* > Hit. *-VIV-* was not at work, the reflex of the base *\*pólH<sub>2</sub>-i-* would be written *\*pa-a-al-hǝ* with the full vowel *a* (the *išpāndi* type).

<sup>12</sup> According to Sturtevant (1933: 106) and Couvreur (1937: 216), an adjective was also derived from the same zero-grade base *\*pǝH<sub>2</sub>-i-* by Oettinger (1979: 550), Melchert (1984: 45; 1994: 55, 125), Kimball (1999: 123, 242), and Wodtko et al. (2008: 562).

<sup>13</sup> The origin of *a*-vocalism in the Hit. alternations *-i-/aǝ-* and *-u-/au-* alongside the IE base *\*-i-/éǝ-* and *\*-u-/éǝ-* is not entirely clear. Melchert (1994: 138) has rejected his earlier attempt to derive this from the *o*-stem variant *\*-i-/óǝ-* and *\*-u-/óǝ-* (Melchert 1984: 45; derived the same way by Kimball 1999: 242, adjective *palhi-* *\*pǝH<sub>2</sub>-i-*, *\*pǝH<sub>2</sub>-óǝ-*) and has decided in favor of an internal Hittite phonetic base for the vowel *a* from PIE *\*e* in post-tonic position and an open syllable. The application of such a phonetic development also makes it possible to explain the present-tense person marker with *a*-vocalism *-yani* (alongside *-yeni*) and *-tani* (alongside *-teni*), where the preterit person markers do not have such vocalism. The transition into an *a*-vowel could have occurred after apocope into *-yēn* < *\*-ye-ne* and *-ten* < *\*-te-ne*. Otherwise, the phenomenon would have also created *a*-vocalism in the preterit person markers.

CLuv. adjective *ṽāšu-* ‘good’, related to Skt. *vasu-* (adj.) and Av. *vohu-* (adj.), as well as the OIr. noun *fó* ‘goodness, obligingness’, points to an *o*-grade < \**ṽosu-*.<sup>14</sup>

The adjectival paradigm from which Hit. *palḫi-* is derived is therefore probably:

nom. sg.	*pólH <sub>2</sub> i-s
acc. sg.	*pólH <sub>2</sub> i-m
nom.-acc. sg.	*pólH <sub>2</sub> i-ø
gen. sg.	*p H <sub>2</sub> -éj̄-s → Hit. <i>palḫi-/palḫaḫ̄-</i> (adj.)

and the nominal paradigm, from which PSl. \**pǝl'e* is derived, is:

nom.-acc. sg. n.	*pólH <sub>2</sub> i-ø → *pólH <sub>2</sub> i-o-m > PSl. * <i>pǝl'e</i> (n.)
gen. sg.	*pélH <sub>2</sub> i-s

4. In the Hittite proterokinetically declined adjective *palḫi-* and the Proto-Slavic noun \**pǝl'e* (n.) it is therefore possible to recognize the old interparadigmatic connection from the earlier protolanguage period, when the *i*-stem and *u*-stem neuter nouns were declined acrostically, but their homophone adjectives proterokinetically. This relationship is all the more valuable because until now such examples in IE languages have been recognized only among *u*-stem formations.

<sup>14</sup> Alongside OIr. *fó* < \**ṽosu* there is also OIr. *feb* ‘goodness, obligingness’ < \**ṽesṽeH<sub>2</sub>*. Alongside this I should draw attention to the same formal relationship between OIr. *serb* ‘theft, robbery’ < \**serṽeH<sub>2</sub>* and Hit. *šāru-* (n.) ‘quarry’ < \**sóru*.

## Bibliography

### References

- BENVENISTE, Emile (1935) *Origines de la formation des noms en indo-européen*. Paris: Librairie d'Amérique et d'Orient Adrien-Maisonneuve.
- COUVREUR, Walter (1937) *De hettitische h. Een bijdrage tot de studie van het indo-europeesche vocalisme*. Leuven: Le Museon.
- FURLAN, Metka (1986) *Indoevropske dvožložne težke baze v hetitščini*. PhD thesis. Ljubljana: University of Ljubljana.
- FURLAN, Metka (1994) "Etimološko raziskovanje slovenskega besedja (Od Jarnikovega do Miklošičevega etimološkega slovarja)." In: M. Orožen (ed) *XXX. seminar slovenskega jezika, literature in kulture*. Ljubljana: Seminar slovenskega jezika, literature in kulture.
- KIMBALL, Sara E. (1999) *Hittite Historical Phonology*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- MELCHERT, H. Craig (1984) *Studies in Hittite Historical Phonology*. Göttingen.
- MELCHERT, H. Craig (1994a) "'Čop's Law' in Common Anatolian." In: J. E. Rasmussen (ed.) *In honorem Holger Pedersen*. Wiesbaden: L. Reichert.
- MELCHERT, H. Craig (1994b) *Anatolian Historical Phonology*. Amsterdam /Atlanta: Rodopi.
- OETTINGER, Norbert (1979) *Die Stammbildung des hethitischen Verbums*. Nürnberg: Hans Carl.
- OLSEN, Birgit Anette (1999) *The Noun in Biblical Armenian. Origin and Word-Formation – with special emphasis on the Indo-European heritage*. Berlin/New York: Mouton de Gruyter.
- PINAULT, Georges-Jean (2003) "Sur les thèmes indo-européen en \*-u- : dérivation et étymologie." In: E. Tichy, D. S. Wodtko, B. Irslinger (eds) *Indogermanisches Nomen. Derivation, Flexion und Ablaut*. Akten der Arbeitstagung der Indogermanischen Gesellschaft. Freiburg, 19. bis 22. September 2001. Bremen: Hempen.
- SCHMITT-BRANDT, Robert (1967) *Die Entwicklung des indogermanischen Vokalsystems (Versuch einer inneren Rekonstruktion)*. Heidelberg: Groos.
- SCHRIJVER, P., 1991: *The reflexes of the Proto-Indo-European Laryngeals in Latin*. Amsterdam/Atlanta: Rodopi.
- STURTEVANT, Edgar H. (1933) *A Comparative Grammar of the Hittite Language*. Philadelphia: Linguistic Society of America.
- WATKINS, Calvert (1982) Notes on the Plural Formation of the Hittite Neuters. In: E. Neu (ed) *Investigationes Philologicae et Comparative, Gedenkschrift für H. Kronasser*. Wiesbaden: L. Reichert.
- WODTKO, Dagmar S./Britta IRSLINGER/Carolin SCHNEIDER (2008) *Nomina im indogermanischen Lexikon*. Heidelberg: Universitätsverlag Winter.



## Dictionaries

- EDSIL: Derksen, Rick *Etymological Dictionary of the Slavic Inherited Lexicon*. Leiden – Boston, 2008.
- EIEC: Mallory, James/Douglas Q. Adams (eds) *Encyclopedia of Indo-European Culture*. London /Chicago: Fitzroy Dearborn Publishers, 1997.
- ESSJ: Bezlaj, France *Etimološki slovar slovenskega jezika I–V*. (Authors of headwords: France Bezlaj, Marko Snoj, Metka Furlan). Ljubljana: Mladinska knjiga, 1976–2007.
- ESJS: *Etymologický slovník jazyka staroslověnskeho II*. Praha: Academia, 2002.
- IEW: Pokorny, Julius *Indogermanisches etymologisches Wörterbuch I*. Bern /München: Francke, 1959.
- HEG: Tischler, Johann *Hethitisches etymologisches Glossar I*. Innsbruck: Institut für Sprachwissenschaft der Universität, 1977.

### Abstract

#### THE ARCHAIC WORD-FORMATIONAL RELATIONSHIP BETWEEN THE PROTO-SLAVIC NOUN \**POL'E* AND THE HITTITE ADJECTIVE *PALHI-*

In the Hittite proterokinetically declined adjective *palhi-* and the Proto-Slavic noun \**pol'e* (n.) it is possible to recognize the old interparadigmatic connection from the earlier protolanguage period, when the *i*-stem and *u*-stem neuter nouns were declined acrostatically, but their homophone adjectives proterokinetically (Watkin's model). This relationship is all the more valuable because until now such examples in IE languages have been recognized only among *u*-stem formations.

### Povzetek

#### ARHAIČNO BESEDOTVORNO RAZMERJE MED PRASLOVANSKIM SAMOSTALNIKOM \**POL'E* IN HETITSKIM PRIDEVNIKOM *PALHI-*

V hetitskem proterokinetično dekliniranem pridevniku *palhi-* in praslovanskem samostalniku \**pol'e* (n.) je mogoče prepoznati staro medparadigmatsko vez iz zgodnejšega prajezičnega obdobja, ko so se ijevski in ujevski samostalniki srednjega spola deklinirali akrostatično, njim homofoni pridevniki pa proterokinetično (Watkinsov model). Razmerje je toliko bolj dragoceno, ker so bili taki primeri do sedaj v ide. jezikih prepoznani le med ujevskimi tvorbami.



## **LA LIAISON DE L'ADJECTIF SUR LE NOM EN FRANÇAIS : MORPHOLOGIE, SYNTAXE, PHONOLOGIE**

Rendre compte de la liaison de l'adjectif sur le nom consistait naguère simplement à décrire les formes liées dans leurs rapports avec les formes libres et à indiquer les classes de contextes dans lesquels apparaissaient ces formes (cf. e.g. Gougenheim 1935 : 89-92). Depuis Schane (1968), la tendance dominante de la linguistique fait au contraire porter le poids de l'explication des faits sur la phonologie : ce serait aux règles de syllabation qu'il revient de déterminer laquelle, de la forme libre ou de la forme liée, doit être employée dans un contexte donné, et même parfois de modeler ces formes en mettant ou non en exercice la consonne finale instable d'une représentation supposée unique. Les pages qui suivent brossent à grands traits ce que pourrait être une conception renouvelée de l'approche traditionnelle. Nous essayons de montrer que les régularités — et les irrégularités — observables dans les formes liées peuvent être expliquées sans qu'on ait à sortir du paradigme flexionnel de l'adjectif, que la sélection des formes ne dépend pas du caractère vocalique ou consonantique du mot qui suit l'adjectif mais du contexte morpho-syntaxique, et que la phonologie ne joue dans cette affaire qu'un rôle assez mince.

L'espace imparti ne nous permet ni de rendre pleine justice à nos prédécesseurs, ni d'essayer de répondre aux objections qui viennent à l'esprit. Nous ne pouvons pas non plus aborder les problèmes que pose le pluriel.

### **1. MORPHOLOGIE**

À la suite de Morin (1992), nous considérons les formes liées du masculin singulier non pas comme des avatars du masculin ordinaire caractérisés par la mise en exercice d'une consonne finale instable devant voyelle, mais comme des membres à part entière du paradigme flexionnel de l'adjectif. Notre position diffère en revanche de celle de Morin (1992, 2003) en ce que nous ne pensons pas que les « consonnes de liaison » soient des marques morphologiques (des suffixes ou des préfixes). À la suite de Boyé et Bonami (2003, 2005, cf. aussi Roché 2010, Boyé et Plénat, à par.), nous opinons que l'adjectif prénominal masculin singulier est un pur radical dont la

---

\* *Adresse de l'auteur*: CLVA, 64 Grand'rue, 81340 Valence d'Albigeois, France. Mél : plenat@univ-tlse2.fr

\*\* *Adresse de l'auteur*: Université de Provence, Technopôle Château-Gombert, 39 rue F. Joliot Curie, 13453 Marseille, France. Mél : plenat@cmi.univ-mrs.fr

forme est déterminée par les rapports réglés qu'il entretient avec les deux thèmes principaux de l'adjectif, le thème A (celui du masculin) et le thème B (celui qu'utilise en particulier le féminin). Ce radical est marqué en ce sens qu'il se termine de préférence par une consonne, mais cette particularité ne confère selon nous aucun statut particulier, ni morphologique, ni phonologique, à cette consonne. Nous ne faisons pas de différence entre le /r/ de *bizarre animal* /bizaranimal/ ou de *lourd anneau* /lurano/ et celui de *dernier étage* /dɛrnjɛrɛtaʒ/ ou /dɛrnjɛrɛtaʒ/. Occupant la même place dans le paradigme, toutes ces formes sont pour nous des formes de liaison du masculin singulier (FLMS).

**1.1. La concurrence entre le thème A et le thème B**

Les FLMS sont normalement apprises. Mis dans l'obligation d'en créer une pour un adjectif donné, les locuteurs sont souvent embarrassés et utilisent des stratégies diverses (cf. Morin, *art. cit.*). Mais ils ne sont pas tous embarrassés au même degré et ne montrent pas le même embarras dans tous les cas. Abstraction faite des cas de supplétion, les FLMS ne recourent qu'au thème A et au thème B. Au sein de cet « espace thématique », les locuteurs peuvent se laisser guider par les deux grandes contraintes partiellement contradictoires qui régissent les FLMS lexicalisées :

- (1) Contrainte 1 : Autant que faire se peut, une FLMS se termine par une consonne ;
- (2) Contrainte 2 : Autant que faire se peut, une FLMS recourt – comme les autres formes du masculin – au thème A.

Il arrive très souvent que, le thème A se terminant par une consonne, ces deux contraintes soient toutes deux immédiatement satisfaites.

	Masc. sing.	FLMS	Fém. sing.
SALE	<i>sale coup</i> /salku/	<i>sale accident</i> /salaksidã/	<i>sale guerre</i> /salger/
VIF	<i>vif succès</i> /vifsykse/	<i>vif intérêt</i> /vifêtere/	<i>vive gratitude</i> /vivgratityd/
COURT	<i>court moment</i> /kurmômã/	<i>court instant</i> /kurêstã/	<i>courte durée</i> /kurtôdyre/

Tableau 1 : FLMS = thème A

Tous les adjectifs dont le thème A se termine par une consonne utilisent aussi ce thème à la FLMS. C'est le cas des adjectifs épiciens comme SALE et des adjectifs qui, comme VIF, comportent au féminin une consonne finale différente de celle du masculin. C'est aussi le cas des adjectifs comme COURT, dont le /r/ est suivi d'une autre consonne au féminin. Il va sans dire que les créations de FLMS nouvelles à partir d'adjectifs dont le masculin s'achève par une consonne opèrent, même dans le dernier cas, suivant ce schéma : il n'est pas besoin d'avoir déjà rencontré *desert* à la liaison pour s'abstenir de faire sonner le *t* dans *un bien desert animal*.

Les choses se compliquent lorsque le masculin se termine par une voyelle. Quand l'adjectif est épïcène (comme par exemple JOLI), le recours exclusif aux éléments de l'espace thématique interdit toute adjonction d'une consonne finale, la FLMS adopte la forme commune aux deux thèmes. Mais quand le thème A se termine par une voyelle et le thème B par une consonne, on observe trois types de comportement :

- 1) La contrainte 1 l'emporte sur la contrainte 2 : autrement dit, le locuteur recourt au thème B (celui du féminin) au lieu de recourir au thème A ;
- 2) Un compromis s'instaure : il est recouru au thème A (celui du masculin), mais la FLMS emprunte au thème B sa ou ses consonnes finales ;
- 3) La contrainte 2 l'emporte sur la contrainte 1 : autrement dit, le locuteur « ne fait pas les liaisons ».

Le premier comportement est considéré comme le plus conservateur. Dans cette parlure, l'emploi du thème B vaut aussi bien pour les adjectifs comme PETIT ou EXACT dont le thème B ne diffère du thème A que par la présence d'une ou de deux consonnes finales, que pour les adjectifs « apophoniques » comme SOT, DERNIER, PLEIN ou NOUVEAU dont le thème B comporte une voyelle différente de celle du thème A.

	Masc. sing.	FLMS	Fém. sing.
PETIT	<i>petit copain</i> /pətikɔpɛ̃/	<i>petit ami</i> /pətitami/	<i>petite copine</i> /pətikɔpin/
EXACT	<i>exact contraire</i> /ɛgzakɔ̃trɛr/	<i>exact opposé</i> /ɛgzaktɔpoze/	<i>exacte mesure</i> /ɛgzaktməzyr/
SOT	<i>sot personnage</i> /sɔpɛrsɔnɑʒ/	<i>sot animal</i> /sɔtanimal/	<i>sotte femme</i> /sɔtfam/
DERNIER	<i>dernier jour</i> /dɛrnjɛʒur/	<i>dernier étage</i> /dɛrnjɛretɑʒ/	<i>dernière main</i> /dɛrnjɛrmɛ̃/
PLEIN	<i>plein soutien</i> /plɛ̃sutjɛ̃/	<i>plein emploi</i> /plɛ̃nɑ̃plwa/	<i>pleine confiance</i> /plɛ̃nkɔ̃fjɑ̃s/
NOUVEAU	<i>nouveau conflit</i> /nuvokɔ̃fli/	<i>nouvel accord</i> /nuvɛlakɔr/	<i>nouvelle guerre</i> /nuvɛlgɛr/

Tableau 2 : FLMS = thème B

Cette stratégie consistant recourir au thème B complet (y compris dans le type EXACT) pour munir la FLMS d'une consonne finale est dans bien des cas facilement généralisable par les locuteurs. L'un de nous, qui n'a très probablement jamais entendu employer *pédant*, *adipeux*, *suspect*, *idiot*, *tracassier*, ou *faubourien* devant un nom masculin singulier commençant par une voyelle, se conformerait assurément à cette stratégie si jamais il avait à lire à haute voix *son pédant ami*, *son adipeux ami*, etc. Il a été surpris le jour où il a appris que l'usage favorisait maintenant l'emploi de /ʒɑ̃ti/ au lieu de /ʒɑ̃tij/ comme FLMS de GENTIL (fém. /ʒɑ̃tij/). Il serait en revanche fort désemparé s'il avait à employer *blanc* dans les mêmes conditions ; ce point sera examiné plus tard.

Le second comportement, considéré comme plus moderne, n'a de conséquences distinctes de celles du premier que pour certains adjectifs apophoniques, car une FLMS comme /pətɪt/ se laisse indifféremment décrire soit comme un thème B, soit comme un thème A suivi de la consonne finale du thème B.

	Masc. sing.	FLMS	Fém. sing.
SOT	<i>sot personnage</i> /sopɛrɔ̃naʒ/	<i>sot animal</i> /sotanɪmal/	<i>sothe femme</i> /sɔ̃tfam/
DERNIER	<i>dernier jour</i> /dɛrnjeʒur/	<i>dernier étage</i> /dɛrnjeʁetaʒ/	<i>dernière main</i> /dɛrnjɛrmɛ̃/
PLEIN	<i>plein soutien</i> /plɛ̃sutjɛ̃/	<i>plein emploi</i> /plɛ̃nɛ̃plwa/	<i>pleine confiance</i> /plɛ̃nkɔ̃fjãs/

Tableau 3 : FLMS = thème A + consonne du thème B

Les adjectifs apophoniques, quant à eux, se laissent ranger en deux catégories : les uns, comme SOT, DERNIER ou PLEIN prennent la voyelle mi-fermée ou nasale du thème A devant la consonne finale du thème B ; en revanche, les adjectifs BEAU, NOUVEAU, FOU, MOU et VIEUX adoptent purement et simplement le thème B dans cette parlure comme dans la parlure traditionnelle : on n'observe jamais de formes comme \*/nuvol/, \*/ful/ ou \*/vjøj/. Il est concevable d'essayer de rendre compte des formes apophoniques en termes de syllabation, comme le fait Féry (2004), mais encore faudrait-il montrer que la consonne de liaison obéit bien à des règles de syllabation différentes de celles de la consonne finale du féminin dans cette parlure. Nous préférons nous ranger à une position proche de celle de Steriade (1999) et expliquer ces formes par un recours simultané à deux formes existant indépendamment dans l'espace thématique de l'adjectif. Comme BEAU et consorts constituent une petite classe fermée, il n'y a pas d'inconvénient à considérer leurs FLMS comme supplétives dans la parlure moderne. Dans les autres cas, le recours simultané à deux formes semble en revanche productif : dans le test de production mené par Féry (2004), *sot* /sot/ et *mignon* /mijɔ̃n/ font des scores plus qu'honorables en dépit du fait que ces formes aient peu de chances de faire partie des FLMS mémorisées par les sujets.

Le troisième comportement est celui des locuteurs les plus novateurs. Tous les spécialistes s'accordent pour estimer que la liaison de l'adjectif sur le nom est en perte de vitesse. Ce phénomène se laisse expliquer sans difficulté par la montée en puissance de la contrainte voulant qu'un adjectif masculin recoure au thème A. Dans les parlures où c'est le cas, les FLMS distinctes de la forme libre deviennent supplétives. Mais le nombre de ces formes supplétives a tendance à se réduire, puisque des formes sans consonne finale se substituent même parfois à des formes bien lexicalisées comme *grand* /grãt/ ou *gros* /groz/.

### 1.2. Formes supplétives et déficiences

Aucune des trois parlures décrites ci-dessus ne se présente jamais sous une forme pure. Chacune d'elles renferme des formes supplétives à côté des formes qui y sont

régulières. On aurait probablement du mal à trouver un locuteur de la parlure novatrice qui n'utilise aucune FLMS spécifique. On a vu que *bel* et consorts devaient être considérés comme supplétifs dans la parlure qui tend à concilier l'exigence de consonne finale et le recours au thème A. Dans le parler traditionnel décrit par Fouché (1959), les adjectifs MALIN et COMMUN conservent leur voyelle nasale à la liaison, le premier au moins dans l'expression ecclésiastique *le malin esprit* /malēnesprit/, le second dans l'expression toute faite *d'un commun accord* /doëkɔmœnakɔr/ et sans doute ailleurs. Fouché signale comme obsolète une FLMS en /k/ qui peut difficilement être tirée du thème B pour FRANC, mais Martinon (1913) pouvait encore dire *un fran(c) kêtourdi*. Il ne mentionne pas, sauf erreur, la FLMS *vieux* /vjøz/, mais celle-ci concurrençait encore *vieil* dans la bonne langue de la première moitié du XX<sup>e</sup> siècle. Cette hétérogénéité ne surprend pas si la majorité des FLMS sont apprises : elles sont empruntées à des traditions diverses.

Il est parfois difficile de décider si une forme est régulière ou supplétive. Le cas se pose notamment dans le cas des FLMS dont la consonne finale est distincte de celle du thème B. On sait que seul un petit nombre de consonnes « de liaison » peuvent figurer à la fin des FLMS. Les thèmes B s'achevant par un /s/ et par un /d/ ne lèguent pas ces consonnes telles quelles aux FLMS : le /s/ se sonorise en /z/ (GROS, fém. *grosse* /gros/, a pour FLMS /groz/ : *gros enfant* /grozãfã/), tandis que le /d/ s'assourdit en /t/ (GRAND, fém. *grande* /grãd/, a pour FLMS /grãt/ : *grand enfant* /grãtãfã/). De même, MALIN, fém. *maligne* /malij/ avec une nasale palatale a pour FLMS /malēn/ avec une dentale dans *malin esprit* /malēnesprit/. Il y a lieu de penser que ces altérations sont d'origine analogique et découlent de l'abondance des cas où la consonne « de liaison » fournie par le thème B est un /z/, un /t/, ou un /n/. Les formes retenues préservent une certaine fidélité à la consonne finale du thème B, tout en cédant à l'attraction de séries nombreuses. (Si la vélaire de la FLMS *long* ne s'assourdit plus guère en /k/ — on prononce *long hiver* /lõgiver/ plutôt que /lõkiver/ —, c'est peut-être qu'il n'existe plus guère dans la langue d'autres cas d'alternance entre /Ø/ et /k/). Mais, quelle qu'en soit l'origine, il n'est pas aisé de déterminer si la sonorisation de /s/ et l'assourdissement de /d/ peuvent être étendus au delà de la petite poignée d'exemples où ils sont bien attestés. Il ne nous semble pas impossible de prononcer un /z/ dans *un épais édreton* ou un /t/ dans *un fécond entretien*, mais tous les locuteurs n'auraient certainement pas la même réaction. Suivant qu'elles sont ou non généralisables, les petites séries de FLMS lexicalisées *bas* /baz/, *doux* /duz/, *faux* /foz/, *gros* /groz/ et *grand* /grãt/, *profond* /prɔfɔt/, *second* /sægɔt/ doivent être considérées comme régulières ou comme supplétives. Les autres adjectifs à féminins en /d/ ou en /s/ sont défectifs pour les locuteurs qu'ils plongent dans l'embarras.

La contrainte qui réduit le nombre des consonnes « de liaison » possibles joue un rôle certain dans la défektivité. En l'absence d'un modèle adéquat dans la langue, les adjectifs dont la consonne finale spécifique au thème B est un /ʃ/ (BLANC, FRAIS, FRANC) sont normalement dépourvus de FLMS. Les locuteurs que l'on contraint à en créer une doivent recourir aux recettes de la parlure novatrice (emploi du thème A : e.g. *franc* /frã/) ou recourir à une consonne finale fréquente

mais sans rapport étroit avec la consonne du thème B (e.g. *franc* /frãt/). Les stratégies d'évitement sont plus naturelles (cf. Morin, *art. cit.*).

Les régularités et les irrégularités constatées plaident les unes et les autres en faveur d'une description des FLMS en termes morphologiques : les supplétions et les déféctuosités montrent que l'on a affaire à un paradigme comportant parfois, comme beaucoup de paradigmes, des cases préremplies et des cases inremplissables ; les régularités suggèrent que les formes qui remplissent les autres cases résultent d'un choix réglé opéré au sein de l'ensemble des thèmes qu'utilisent le masculin et le féminin.

## 2. SYNTAXE

Cette description morphologique de la forme que prennent les FLMS n'exclut pas *a priori* que le choix de ces formes réponde à la nécessité de fournir une attaque initiale au nom-tête ou de prévenir un hiatus dans le contexte prénominal. Mais elle ouvre aussi la voie à une approche dans laquelle la syntaxe prend une part prépondérante dans le conditionnement de ce choix. La présence de la consonne finale des FLMS étant traitée comme une affaire interne au paradigme de l'adjectif, la possibilité s'ouvre que la sélection de la classe des FLMS soit une sorte de phénomène d'accord imposé par la présence d'un nom appartenant à une classe de déclencheurs de liaison. Deux arguments militent en ce sens : en premier lieu, l'existence de phénomènes de « liaison à distance » au sein du groupe nominal, en second lieu celle de liaisons non-enchaînées.

### 2.1. La liaison à distance dans le groupe nominal

Parlant au nom de l'Office de la langue française, Dauzat (1938) signalait que les adjectifs BEAU et NOUVEAU prennent à l'occasion la forme de *bel* et de *nouvel* quand ils constituent le premier membre d'une coordination adjectivale antéposée à un nom à initiale vocalique. Selon lui, on peut dire *un bel et charmant enfant, son nouvel et dernier avatar*, alors que, lorsque le nom commence par une consonne, on est forcé d'employer la variante antéconsonantique. Placée à distance de la voyelle qui est censée déclencher sa présence, la FLMS ne peut guère, dans ce cas, être choisie pour fournir une attaque à la syllabe initiale d'un nom auquel, au demeurant, le second adjectif en fournit déjà une. Son choix ne répond pas non plus à la nécessité d'éviter un hiatus avec la conjonction *et*, puisque l'hiatus est de règle quand le nom-tête commence par une consonne (*un beau et charmant garçon*) ou que la coordination adjectivale est postposée (*un enfant beau et charmant*). En revanche, cette situation est attendue dans une solution morpho-syntaxique : deux adjectifs coordonnés ayant normalement les mêmes flexions, si l'un est à la FLMS, l'autre doit l'être aussi. Dauzat ne fondait sa remarque que sur le sentiment de cinq membres de l'Office, qui n'étaient pas d'accord entre eux, et considérait l'emploi de *bel* et *nouvel* dans cette construction comme une tournure traditionaliste, la langue moderne préférant selon lui *beau* et *nouveau*. Qu'en est-il de nos jours ?



Nous avons (Plénat, à par.) examiné les coordinations adjectivales en *et* antéposées dont le premier terme est BEAU, NOUVEAU ou l'un des trois autres adjectifs (FOU, MOU, VIEUX) dont la FLMS a une orthographe distincte de celle du masculin ordinaire d'une part dans dix années du *Monde* (1991-2000), et, d'autre part, sur la Toile (en nous restreignant dans ce dernier cas aux coordinations précédées de l'article défini). Le résultat est surprenant. Dans l'un et l'autre corpus, la FLMS est employée dans sensiblement plus de 90% des cas où le nom-tête fait partie de la classe des noms déclencheurs de liaison, alors que, lorsque le nom inhibe les liaisons, cette forme n'apparaît que dans environ 5% des cas dans *Le Monde* et dans environ 15% des cas sur la Toile. Les chiffres exacts varient légèrement en fonction de la façon dont on définit la classe des déclencheurs de liaison et doit tenir compte de cas particuliers comme l'expression figée *bel et bon*, nous épargnons ici ces détails au lecteur. Quels que soient les pourcentages exacts, il apparaît clairement qu'il est de nos jours de règle d'employer ces FLMS dans les coordinations adjectivales en *et* antéposées à un nom déclencheur de liaison. À titre d'illustration, on trouvera en (1) des exemples trouvés sur la Toile où les FLMS sont précédées de *son* :

(1) *son bel et puissant engin ; son nouvel et modeste appartement ; son fol et aveugle espoir ; son mol et impérial époux ; son vieil et riche amant*

Les coordinations de deux adjectifs ne sont pas les seules à autoriser la présence d'une FLMS en première position devant un nom déclencheur de liaison. Nous n'avons pas examiné systématiquement les cas où une telle coordination est juxtaposée à un troisième adjectif, et nous n'avons donc pas de pourcentages à proposer, mais on trouve bien sur la Toile des exemples où ce troisième adjectif prend la forme d'une FLMS, cf. :

(2) *un bel, intéressant et rare ensemble ; nouvel, utile et extraordinaire écosystème ; un vieil, aimable et excellent homme*

Le problème se pose de savoir si le comportement qui vient d'être décrit est le fait des seuls adjectifs BEAU, NOUVEAU, FOU, MOU et VIEUX, ou si tous les adjectifs apparaissent sous la forme de leur FLMS dans le contexte étudié. L'orthographe n'est ici d'aucun secours, puisqu'en général la graphie de la FLMS n'est pas distincte de celle de la forme libre. On en est réduit à l'intuition, ... et aux graphies fautives. Pour ce qui est de la première, il ne fait pas de doute pour nous que certains adjectifs se comportent comme BEAU et consorts. Nous dirions *un charmant t-et bel enfant* comme nous disons *un bel et charmant t-enfant*, ou *un grand t-et bel appartement* comme *un bel et grand t-appartement*. Il est probable que nous ne ferions pas la liaison dans tous les cas, et que les locuteurs varient sur ce point. Seule une batterie de tests sur un panel de sujets parlants permettrait de déterminer les facteurs qui favorisent ou inhibent la liaison dans ces contextes très rares. Mais il ne fait pas de doute que des liaisons apparaissent avec d'autres adjectifs que BEAU et consorts.

Cette existence est confirmée par les fautes d'orthographe. Certains scripteurs écrivent les FLMS comme les féminins avec lesquels ils se confondent phonétiquement. Le fait est courant notamment dans les formules de vœux d'anniversaire, qui empruntent au style noble des coordinations redondantes comme BON *et* HEUREUX, JOYEUX *et* HEUREUX, etc. Nous avons trouvé sur la Toile environ 150 exemples comme les suivants :

(3) *je te souhaite en cet journée un bonne et heureux anniversaire que cet année t'apporte chance, amour, bonheur et argent rien que du bon quoi*  
*coucou titicia c méliisa ba ct pr te souhaiter un joyeuse et heureuse anniversaire sa fai koi davoïr 18 ans allez bisou ma poule*

C'est là un tout petit nombre dans l'Océan de la Toile. Mais, parmi les vœux de Noël, pourtant beaucoup plus nombreux que les vœux d'anniversaire, nous n'avons trouvé qu'une occurrence de *un bonne et heureux Noël* avec la graphie féminine de BON. Il ne fait pas de doute que le nombre beaucoup plus élevé des graphies féminines avec *anniversaire* est lié au fait que ce mot est un déclencheur de liaison, et que *bonne* ou *joyeuse* sont dans ces expressions des graphies fautives de FLMS.

Les adjectifs qualificatifs ne sont pas seuls à être sensibles à ce type de conditionnement à distance. Schlenker (à par.) signale un comportement analogue des possessifs : dans les constructions où un premier possessif d'une personne du singulier est coordonné à un second possessif, le premier possessif apparaît sous la forme de *mon, ton, son* et non de *ma, ta, sa* lorsque les possessifs sont suivis d'un syntagme féminin déclencheur de liaison :

(4) *Priscilla a été successivement mon (/mõ/), et ton (/tõn/) épouse vs Priscilla a été successivement ma, et ta femme*

Schlenker note qu'il y a dans ces constructions un phénomène de supplétion plutôt que de liaison : dans l'exemple (4), *mon* prend sa forme antéconsonantique /mõ/ et non sa forme antévocalique /mõn/. Cette discordance complique l'analyse des faits. Mais ceux-ci montrent bien que les deux termes d'une coordination relèvent de la même classe d'allomorphes.

Nous aurons l'occasion de revenir ci-dessous sur les coordinations adjectivales pour montrer que l'emploi des FLMS y obéit à un conditionnement un peu plus complexe que ce qui vient d'être dit. Mais le caractère massif de cet emploi lorsque le nom-tête est un déclencheur de liaison constitue un bon argument en faveur du caractère syntaxique de la liaison de l'adjectif masculin singulier sur le nom.

## 2.2. La liaison sans enchaînement

Lorsque la seconde FLMS d'une coordination adjectivale prend place dans une incise, il arrive que la consonne finale de l'adjectif s'enchaîne sur le nom malgré la

rupture intonative et l'éventuelle pausule qui séparent les deux mots : on dit *un robuste, mais petit, t-enfant* (Pichon 1938). Ce comportement contraste avec celui du féminin, dont la consonne finale n'est jamais enchaînée : *une robuste, mais petite, enfant*. Cette particularité des consonnes finales des FLMS du type de *petit* a apporté de l'eau au moulin des tenants de la thèse voulant que les consonnes finales « instables » jouissent d'un statut – phonologique ou morphologique – particulier (cf. e.g. Morin 1986 : 198-199). Comment, en effet, expliquer par exemple que *petit* et *petite* aient deux comportements phonologiques distincts s'ils ont une représentation phonologique commune ? Notre réponse (cf. Plénat 2008) consiste à dire que les enchaînements après une rupture intonative ne sont qu'une conséquence accessoire du caractère partiellement contradictoire des contraintes pesant sur la sélection des formes libres et des formes liées dans ce type de contexte.

Le comportement des adjectifs devant une rupture intonative a été étudié principalement dans une construction « disloquée » dans laquelle le nom fournit une référence à un pronom qui le précède, comme dans *j'en ai un petit, t-éléphant* (cf. Tranel 1990, Steriade 1999, Féry 2003). C'est cette construction que nous utiliserons ici. Sa grammaticalité est douteuse, dans la mesure où le nom est privé de tout déterminant ; elle résulte en quelque sorte de la superposition d'une phrase complète (*J'en ai un petit*) et d'un SN (*un petit éléphant*). Mais les intuitions phonologiques auxquelles elle donne lieu paraissent assez sûres, même pour les locuteurs qui la rejettent comme agrammaticale.

Tous les auteurs semblent d'accord pour dire que l'enchaînement n'a lieu que lorsque la forme qu'il laisse devant la pausule est identique à la forme libre du masculin. C'est là le cas des FLMS empruntées au thème B (type PETIT) ou inspirées du thème B (type GRAND) – à partir du moment du moins où leur voyelle est identique à celle du thème A, et de celles qui ne se laissent décrire que comme des thèmes A augmentés de la consonne finale du thème B (comme dans *sot /sot/*) ou d'une autre consonne (comme dans *vieux /vjøz/*).

<i>J'en ai un petit, éléphant</i>	ʒãneẽpti    telefã	cf. il est petit	ilepti
<i>J'en ai un grand, éléphant</i>	ʒãneẽgrã    telefã	cf. il est grand	ilegrã
<i>J'en ai un sot, éléphant</i>	ʒãneẽso    telefã	cf. il est sot	ileso
<i>J'en ai un vieux, éléphant</i>	ʒãneẽvjø    zelefã	cf. il est vieux	ilẽvjø

Tableau 4 : Liaisons enchaînées

Quand, en revanche, la FLMS ne se laisse pas décrire comme une forme libre suivie d'une consonne, l'enchaînement devient impossible. C'est là le cas des adjectifs dont le thème A se termine par une consonne (comme SALE, VIF ou COURT) et des adjectifs dont la FLMS est identique à un thème B à voyelle apophonique (type NOUVEAU), même dans le cas où ce n'est là qu'une variante (comme dans *sot /sɔt/*).

<i>J'en ai un sale, éléphant</i>	ʒãñeẽsal    elefã	cf. il est sale	ilesal
<i>J'en ai un disert, éléphant</i>	ʒãñeẽdizer    elefã	cf. il est disert	iledizer
<i>J'en ai un bel, éléphant</i>	ʒãñeẽbel    elefã	cf. il est beau	ilebo
<i>J'en ai un sot, éléphant</i>	ʒãñeẽsot    elefã	cf. il est sot	ilesos

Tableau 5 : Liaisons non enchaînées

Même des locuteurs qui n'emploient pas habituellement une FLMS donnée se sentent forcés de se conformer cette généralisation si on leur demande d'employer cette FLMS. Nous dirions naturellement, sans enchaînement, /ʒãñeẽsot || elefã/ pour *j'en ai un sot, éléphant* et /ʒãñeẽvjɛj || elefã/ si l'animal est vieux ; mais si nous nous obligions à employer les FLMS /sot/ et /vjøz/, nous dirions à coup sûr, en enchaînant, /ʒãñeẽso || telefã/ et /ʒãñeẽvjø || zelefã/.

Notre hypothèse repose sur l'idée que, dans ces constructions, l'adjectif doit apparaître à la fois sous sa forme libre en tant qu'élément final d'une proposition et sous la forme d'une FLMS en tant qu'adjectif prénominal. Les deux contraintes peuvent être satisfaites simultanément. C'est ce qui se passe lorsque les deux formes sont indistinctes ou lorsque la FLMS se laisse décrire comme une forme libre suivies d'une consonne. Dans le premier cas, l'enchaînement ne peut pas avoir lieu, sous peine d'avoir devant la rupture intonative une forme distincte de tous les allomorphes de l'adjectif (ex. \**j'en ai un sa* || *l-éléphant*) ; dans le second, l'enchaînement est nécessaire, sous peine que ce ne soit pas la forme libre qui apparaisse devant la rupture intonative (ex. \**j'en ai grand t-* || *éléphant*). Mais quand la FLMS ne se confond pas avec le thème A et n'inclut pas strictement ce thème, les deux contraintes ne peuvent pas être satisfaites simultanément, et c'est alors la contrainte pesant sur l'adjectif prénominal qui prend le pas.

Les quantificateurs montrent bien, à notre sens, que cette solution à des chances d'être la bonne. Le paradigme de ces éléments comporte trois cases : celle de déterminant antévocalique (ex. *plus* /plyz/, *dix* /diz/), celle de déterminant antéconsonantique (ex. *plus* /ply/, *dix* /di/) et la forme libre, qui sert dans les autres cas (ex. *plus* /plys/, *dix* /dis/). Dans certains cas, la forme antévocalique ou la forme antéconsonantique se confondent avec la forme libre (ex. *huit* : f. antévoc. et f. libre /ɥit/, f. antécons. /ɥi/ ; *moins* : f. antévoc. /mwẽz/, f. antécons. et f. libre /mwẽ/ ; *deux* : f. antévoc. /døz/, f. antécons. et f. libre /dø/).

<i>Il est plus, aimable</i>	illeplys    emabl	<i>Il est plus</i>	illeplys
<i>Il est moins, aimable</i>	illemwẽ    zemabl	<i>Il est moins</i>	illemwẽ
<i>J'en ai dix, éléphants</i>	ʒãñedis    elefã	<i>J'en ai dix</i>	ʒãñedis
<i>J'en ai huit, éléphants</i>	ʒãñeɥit    elefã	<i>J'en ai huit</i>	ʒãñeɥit
<i>J'en ai deux, éléphants</i>	ʒãñedø    zelefã	<i>J'en ai deux</i>	ʒãñedø

Tableau 6 : Les quantificateurs

Dans les constructions « disloquées », c'est toujours la forme libre (/plys/, /mwẽ/, /dis/, /qit/, /dø/) qui figure devant la rupture intonative, même lorsque cette forme est distincte des deux formes de déterminant (cas de /plys/ et de /dis/). Cet aspect de la distribution des formes montre que c'est le contexte syntaxique de fin de phrase qui dicte le choix de la forme : /plys/ et /dis/ sont normalement incompatibles avec la position de déterminant. Toutefois, lorsque la forme libre est incluse strictement dans le déterminant antévocalique (cas de /mwẽ-z/ et de /dø-z/) et que le nom est un déclencheur de liaison, la consonne finale du déterminant apparaît, mais enchaînée au nom après la rupture intonative. Ce qui montre que le quantificateur est analysé simultanément comme un déterminant antévocalique : la consonne /z/ est incompatible avec les autres contextes. Dans le cas de *huit*, dont la forme libre et la forme antévocalique sont indistinctes, les deux contraintes sont satisfaites simultanément sans qu'il soit besoin d'enchaîner la consonne finale sur le nom.

Le comportement des quantificateurs dans les constructions disloquées ne laisse pas de doute sur le fait qu'une contrainte syntaxique impose que la forme apparaissant devant la rupture intonative soit une forme libre. Dans leur cas, cette contrainte s'impose dans tous les cas de figure. En ce qui concerne les adjectifs, c'est la contrainte imposant la forme liée qui l'emporte toujours, mais il n'y a pas lieu pour autant de penser que la contrainte imposant la forme libre en fin de proposition n'est pas active. Dans l'un et l'autre cas, l'enchaînement n'est qu'une conséquence accessoire d'un conflit opposant deux contraintes syntaxiques, une solution de compromis, et non l'indice d'un statut spécial des consonnes « instables ».

### 3. PHONOLOGIE

La phonologie intervient dans la définition de la classe des noms déclencheurs de liaison, et de la classe complémentaire, mais indirectement. Les mots commençant par une consonne appartiennent tous à la classe des inhibiteurs de liaison, mais c'est, comme on sait, une simplification abusive que d'assimiler la classe de déclencheurs de liaison à celle des noms à initiale vocalique. Les mots « en *h* aspiré » commencent par une voyelle sans pour autant autoriser la liaison, et ce n'est pas une mince affaire que de cerner l'ensemble qu'ils forment (cf. Cornulier 1981). De même, les mots commençant par ces attaques potentielles que sont les semi-voyelles n'ont pas un comportement uniforme. Un traitement phonologique de ces problèmes doit utiliser des éléments abstraits pour différencier les lexèmes qui autorisent la liaison de ceux qui l'interdisent. Une approche morpho-syntaxique utilisant des diacritiques pour distinguer les lexèmes déclencheurs de liaison des inhibiteurs de liaison ne laisse pas échapper de généralisations importantes. Le choix des FLMS peut donc bien être traité fondamentalement comme un phénomène de compatibilité entre classes morpho-syntaxiques.

Il n'est probablement pour autant pas vrai que la phonologie ne joue aucun rôle dans la répartition des FLMS. Nous avons repéré (cf. Plénat, à par.), au sein des coordinations d'adjectifs, deux types de contextes phonologiques qui favorisent le premier l'emploi des formes libres, le second celui des FLMS.

### 3.1. Les coordinations adjectivales en *mais*

Si, dans les coordinations adjectivales, le choix des variantes était déterminé exclusivement par la classe du nom-tête, la présence d'une consonne après le premier adjectif ne devrait avoir aucune influence. Ce n'est pas tout à fait le cas.

Nous avons trouvé sur la Toile un peu plus d'une centaine d'exemples des FLMS *bel*, *nouvel* et *vieil* utilisées comme premier terme d'une coordination en *mais* antéposée. Dans 99% des cas, le nom-tête du syntagme nominal appartenait à la classe des déclencheurs de liaison (cf. exemples (5)). Une recherche moins systématique nous a aussi fait découvrir quelques cas où ces formes précédaient une coordination adjectivale antéposée dont le premier adjectif commençait par une consonne (cf. exemples (6)). Le nom-tête de ces coordinations appartenait à la classe des déclencheurs de liaison dans la majorité des cas.

(5) *un bel (mais froid) après-midi de novembre ; ce bel mais fragile ordonnancement ; un nouvel, mais hélas éphémère, âge d'or ; un nouvel mais bref élan national ; le vieil (mais imparable) adage ; mon vieil mais toujours ronronnant Athlon XP 2000+ argentine*

(6) *ce bel, grand et fière [sic] étalon noir ; ce nouvel, bel et vaste espace ; ce vieil, grand et bel arbre*

Ces données sont on ne peut plus propres à nous conforter dans l'idée que, dans les coordinations adjectivales antéposées, les FLMS sont sélectionnées non pas pour fournir une attaque au mot suivant ou pour éviter un hiatus, mais bien en fonction de la classe du nom-tête.

Il existe cependant des différences importantes entre les coordinations en *et* et les coordinations en *mais*. En premier lieu, lorsque la conjonction est *mais*, les FLMS *bel*, *nouvel* et *vieil* sont proportionnellement moins nombreuses, à quelque classe qu'appartienne le nom-tête. Nous venons de voir que, sur la Toile, ces formes n'étaient employées avec *mais* que dans 1% des cas lorsque le nom-tête était un inhibiteur de liaison ; nous avons signalé auparavant (§ 2.1), que dans le même contexte, ces FLMS apparaissaient dans environ 15% des cas lorsque la conjonction est *et*. Nous avons vu aussi que, lorsque le nom-tête était un déclencheur de liaison, *bel* et consorts apparaissaient dans plus de 90% des cas avec *et* ; avec *mais*, ce pourcentage est à peine supérieur à 60%, et tombe à 50% si l'on fait abstraction des cas, relativement nombreux, où la ponctuation suggère que le second terme de la coordination est une incise. En second lieu, il nous paraît plus que douteux que tous les adjectifs puissent facilement figurer sous la forme de leur FLMS dans les coordinations en *mais* précédant un déclencheur de liaison. Nous ne disposons pas de données écrites sur ce point. Mais il nous semble par exemple difficile de dire *un dangereux z- mais bel animal* ou *un élégant t-, mais vieil, uniforme*, alors que *un vieil, mais élégant, t-uniforme* ou *un bel mais dangereux z-animal* nous paraissent parfaits. Nos intuitions sont flottantes dans beaucoup de cas, mais comme la première, cette seconde différence plaide en faveur de l'idée que la phonologie joue un rôle dans

la sélection des variantes adjectivales. Plus précisément, il semble que le fait que, devant la consonne initiale de la conjonction, la consonne finale de l'adjectif soit amenée à jouer le rôle d'une coda – et d'une coda marquée dans le cas des obstruantes – soit de nature à neutraliser l'influence du nom-tête. Ce point mériterait des enquêtes supplémentaires.

### 3.2. Le rôle du second adjectif

Nous avons mentionné ci-dessus (§ 2.1) le fait que, dans une minorité de cas (5% dans *Le Monde*, 15% sur la Toile), on trouvait des FLMS dans des coordinations adjectivales en *et* préposées à un nom inhibiteur de liaison. Voici des exemples trouvés dans *Le Monde* :

(7) *le bel et austère château de Charles Quint ; le nouvel et audacieux pari d'Edgar Morin ; ce vieil et incorrigible « peacenik »*

Ces exceptions à la règle commune semblent ne pas se répartir aléatoirement. Elles sont, proportionnellement, beaucoup plus nombreuses lorsque le second adjectif commence par une voyelle, comme dans les exemples ci-dessus (cf. *austère, audacieux, incorrigible*). Dans *Le Monde* – où nous n'avons trouvé à vrai dire qu'une poignée d'exemples –, l'écart est de 1 à 6 ; dans les données de la Toile que nous avons recueillies systématiquement (constructions commençant par un article défini), l'écart n'est que de 1 à 3, mais les données sont assez nombreuses pour qu'on puisse avoir la conviction que cette répartition n'est pas le fruit du hasard. Nous avons aussi procédé à des sondages sur la répartition de coordinations adjectivales de sens voisin ne s'opposant que par la nature vocalique ou consonantique de l'initiale du second adjectif comme *nouvel et ultime* et *nouvel et dernier*. Comme on s'y attend, ces coordinations apparaissent les unes et les autres très majoritairement devant des noms déclencheurs de liaison comme *album*. Mais les cas où elles précèdent un autre nom sont beaucoup plus nombreux lorsque le second adjectif commence par une voyelle que lorsqu'il commence par une consonne. *Nouvel et dernier*, par exemple, précède un nom à initiale consonantique comme *disque* dans moins de 1% des cas, alors que *nouvel et ultime* apparaît dans ce type de contexte dans près de 20% des cas. Nous attribuons cet emploi des FLMS non pas au fait que, lorsqu'il commence par une voyelle, le second adjectif serait un déclencheur de liaison, mais au fait que la consonne finale de la FLMS permet d'éviter l'apparition d'un double hiatus. Corroborent cette interprétation d'une part le fait que les adjectifs « en *h* aspiré » paraissent se comporter comme les autres adjectifs à initiale vocalique, et, d'autre part, le fait que le phonème initial du second adjectif semble n'avoir aucune influence dans les coordinations en *mais*. Prenez par exemple le cas de *énième*, qui, pour la plupart des locuteurs, commence par un « *h* aspiré ». Dans nos données, la coordination *nouvel et énième* apparaît dans près de 25% des cas devant un nom à initiale consonantique. Ce comportement, très analogue à celui de *nouvel et ultime*, ne peut être attribué qu'au fait que *énième*, comme *ultime*, commence par une voyelle

et fait donc hiatus avec la conjonction *et*. L'emploi de *nouveau* au lieu de *nouvel* provoquerait l'apparition d'un double hiatus (/nuvoeenjɛm/, /nuvoeyltim/). Commencant par une consonne, la conjonction *mais* empêche l'apparition d'un double hiatus quand le second adjectif commence par une voyelle, et, dans nos données, l'emploi des FLMS devant *mais* est à peine supérieur à 1% que le second adjectif commence par une voyelle ou par une consonne.

Ces influences du contexte proprement phonologique sur le choix entre FLMS et forme libre devront faire l'objet d'enquêtes plus étendues et d'analyses plus élaborées. Le fait le plus digne d'attention, c'est que des FLMS peuvent apparaître devant consonne si le contexte syntaxique s'y prête.

#### 4. CONCLUSION

Il ne serait probablement pas utile de rappeler ici nos conclusions principales. Le lecteur les retrouvera au besoin condensées dans le résumé qui accompagne cet article. Qu'il nous soit simplement permis d'insister sur le rôle de faits statistiquement marginaux comme la liaison à distance ou la liaison enchaînée après rupture intonative. Ces particularités restreignent considérablement la gamme – très étendue – des interprétations possibles. C'est de ce côté, pensons-nous, qu'il conviendrait de rassembler des données plus nombreuses et d'affiner les analyses.

#### Références

- ANDERSEN, Henning (ed.) (1986) *Sandhi Phenomena in the Languages of Europe*. Berlin/New York/Amsterdam : Mouton de Gruyter.
- BONAMI, Olivier/Gilles BOYÉ (2003) « La nature morphologique des allomorphies conditionnées : les formes de liaison des adjectifs en français. » In : B. Fradin/G. Dal/N. Hathout/F. Kerleroux/M. Plénat/M. Roché (éds), 39-48.
- BONAMI, Olivier/Gilles BOYÉ (2005) « Construire le paradigme d'un adjectif. » *Recherches linguistiques de Vincennes* 34, 77-98.
- BOYÉ, Gilles/Marc PLÉNAT (à paraître) « L'allomorphie radicale dans les lexèmes adjectivaux du français. Le cas des adverbes en *-ment*. » In : Tranel (ed.).
- BULLOCK, Barbara E./Jean-Marc AUTHIER/Lisa A. REED (eds.) (1999) *Formal Perspectives in Romance Linguistics*. Amsterdam : John Benjamins.
- CLAS, André (éd.) (1992) *Le mot, les mots, les bons mots*. Montréal : Presses de l'Université de Montréal.
- CORNULIER, Benoît de (1981) « H-aspiré et la syllabation : expressions disjonctives. » In : D. Goyvaerts (ed.), 183-230.
- DAUZAT, Albert (1938) « Autour de l'Office de la langue française. » *Le Français moderne* 6-3, 212.
- DURAND, Jacques/Benoît HABERT/Bernard LAKS (éds) *Congrès Mondial de Linguistique Française – CMLF'08*. Paris : EDP Sciences [<http://www.linguistiquefrancaise.org/>].



- FÉRY, Caroline (2004) « Liaison and syllable structure in French. » Manuscrit, Potsdam.
- FOUCHÉ, Pierre (1959) *Traité de prononciation française*. Paris : Klincksieck.
- FRADIN, Bernard/Georgette DAL/Nabil HATHOUT/Françoise KERLEROUX/Marc PLÉNAT/Michel ROCHÉ (éds) *Silexicales 3. Les unités morphologiques*. Lille : SILEX, Université de Lille 3.
- GOUGENHEIM, Georges (1935) *Éléments de phonologie française*. Strasbourg : Publications de la Faculté des lettres de Strasbourg.
- GOYVAERTS, Didier (ed.) (1981) *Phonology in the 1980's*. Ghent : Story-Scientia.
- MARTINON, Philippe (1913) *Comment on prononce le français*. Paris : Larousse.
- MORIN, Yves-Charles (1986) « On the morphologization of word-final consonant deletion in French. » In : H. Andersen (ed.), 167-210.
- MORIN, Yves-Charles (1992) « Un cas méconnu de la déclinaison de l'adjectif français : les formes de liaison de l'adjectif antéposé. » In : A. Clas (éd.), 233-250.
- MORIN, Yves-Charles (2003) « Remarks on prenominal liaison consonants in French. » In : S. Ploch (ed.), 385-400.
- PICHON, Edouard (1938) « Genre et questions connexes (sur les pas de M<sup>elle</sup> Durand). » *Le Français moderne* 6, 29-43 et 107-126.
- PLÉNAT, Marc (2008) « La liaison « obligatoire » avec et sans enchaînement. » In : J. Durand/B. Habert/B. Laks (éds), 1657-1667.
- PLÉNAT Marc (à paraître) « La liaison à distance dans le groupe nominal. » *Le Français moderne*.
- PLOCH, Stefan (ed.) *Living on the Edge. 28 Papers in Honour of Jonathan Kaye*. Berlin/New York : Mouton de Gruyter.
- ROCHÉ, Michel (2010) « Base, thème, radical. » *Recherches linguistiques de Vincennes* 39, 95-133.
- SCHANE, Sanford A. (1968) *French Phonology and Morphology*. Cambridge, Mass. : MIT Press.
- SCHLENKER, Philippe (à paraître) « A Phonological Condition that Targets Discontinuous Syntactic Units: *ma/mon* suppletion in French ». *Snippets*.
- STERIADE, Donca (1999) « Lexical Conservatism and French Adjectival Liaison. » In : B. Bullock/J.-M. Authier/L. Reed (eds.), 243-270.
- TRANDEL, Bernard (1990) « On Suppletion and French Liaison. » *Probus* 2, 169-208.
- TRANDEL, Bernard (ed.) (à paraître). *Understanding Allomorphy. Perspectives from Optimality Theory*. Londres : Equinox Publishing.

## Résumé

### LA LIAISON DE L'ADJECTIF SUR LE NOM EN FRANÇAIS : MORPHOLOGIE, SYNTAXE, PHONOLOGIE

La liaison de l'adjectif sur le nom en français relève pour l'essentiel de la morphologie et de la syntaxe. Seules étudiées ici, les formes de liaison du masculin singulier (FLMS) constituent l'un des éléments du paradigme flexionnel de l'adjectif. Un adjectif peut être défectif à la FLMS ou prendre une FLMS supplétive. Les formes régulières, quant à elles, résultent de l'interaction de deux contraintes partiellement contradictoires : une FLMS doit, autant que faire se peut, 1) recourir au thème utilisé au masculin et 2) se terminer par une consonne. D'où, en cas de conflit, diverses stratégies consistant à retenir le thème utilisé au féminin (ex. SOT, FLMS : /sɔt/), à privilégier le thème du masculin (ex. SOT, FLMS : /so/), ou à munir ce dernier de la consonne finale du thème du féminin (ex. SOT, FLMS : /sot/). Les FLMS sont choisies non pas en vue de fournir une attaque à la syllabe initiale du nom qui suit, mais simplement sélectionnées par une classe de noms déclencheurs de liaison qui ne se laisse pas définir en termes purement phonologiques. Les « liaisons à distance » que l'on observe lorsque deux adjectifs coordonnés sont préposés au nom (cf. *un bel et charmant enfant*), même quand la conjonction commence par une consonne (cf. *un bel, mais chétif, enfant*) montrent suffisamment que la sélection des FLMS est une opération syntaxique et non phonologique. Ce sont aussi des contraintes syntaxiques qui expliquent la répartition des cas d'enchaînement et de non-enchaînement lorsque l'adjectif est séparé du nom par une rupture intonative et une pause (cf. *j'en ai un petit, t-éléphant* vs *j'en ai un bel, éléphant*). Dans ce contexte, l'adjectif devrait apparaître à la fois sous sa forme liée en tant qu'il est prénominal et sous sa forme libre devant la rupture intonative. Dans tous les cas, c'est la FLMS qui est sélectionnée, mais, lorsque celle-ci inclut strictement la forme libre (cas de /pətɪt/, forme libre /pəti/), l'enchaînement permet de satisfaire en même temps la contrainte requérant la présence de la forme libre devant la rupture intonative. Ces enchaînements ne sont en aucun cas l'indice d'un statut spécial – phonologique ou morphologique – des « consonnes de liaison ». Curieusement, on constate toutefois qu'au sein des coordinations adjectivales, la phonologie joue à l'occasion un rôle dans la sélection des FLMS : celles-ci sont proportionnellement moins nombreuses devant consonne que devant voyelle, mais plus nombreuses quand leur présence permet d'éviter un double hiatus.

VEZANJE PRIDEVNIKA S SAMOSTALNIKOM V FRANCOŠČINI :  
MORFOLOGIJA, SINTAKSA, FONOLOGIJA

Vežanje pridevnika s samostalnikom v francoščini sodi večinoma na področje morfologije in sintakse. V pričujočem članku se bomo ukvarjali izključno z oblikami vežanja v moškem spolu (OVMS) kot enem od elementov pregibne paradigme pridevnika. Oblika vežanja pridevnika je lahko ali neobičajna ali pa nadomestna. Pravilne oblike izvirajo iz učinkovanja dveh delno nasprotujočih si silnic : OVMS se mora v največji mogoči meri (1) opreti na osnovo moške oblike in (2) se končati s soglasnikom. Od tu izvirajo v primeru razhajanja različne strategije, ki poskušajo ali ohraniti osnovo v ženskem spolu (npr. *sot*, OVMS : /sot/) ali dati prednost moški osnovi (npr. *sot*, OVMS : /so/) ali pa moški osnovi dodati končni soglasnik ženske osnove (npr. *sot*, OVMS : /sot/). OVMS niso izbrane, da bi služile kot vzglasje začetnemu zlogu samostalnika, ki sledi, temveč so izbrane glede pripadnosti posamezni vrsti samostalnikov glede na vežanje, ki ga vpeljejo, ki pa ga ni mogoče opredeliti zgolj na ravni fonologije. »Oddaljeno vežanje«, do katerega pride, ko sta dva priredno povezana pridevnika postavljena pred samostalnik (npr. *un bel et charmant enfant*), tudi tedaj ko se vezniška beseda začne s soglasnikom (npr. *un bel, mais chétif, enfant*), nazorno dokazuje, da je izbira OVMS sintaktično in ne fonološko pogojena. Prav tako je mogoče samo na sintaktični ravni pojasniti primere navezovanja oziroma nenavezovanja, ko intonacijski prelom oziroma krajši premor pridevnik loči od samostalnika (primerjaj: *j'en ai un petit, t-éléphant* proti *j'en ai un bel, éléphant*). V takšnem sobesedilnem okolju bi se moral pridevnik pojaviti tako v vežani obliki, ko stoji pred samostalnikom, kot v nevezani, ko stoji pred intonacijskim prelomom. V vseh primerih je izbrana OVMS, a ko ta vključuje nevezano obliko (primer pridevnika /pətit/, nevezana oblika /pəti/) tedaj lahko z navezovanjem zadostimo tudi zahtevi po prisotnosti nevezane oblike pred intonacijskim prelomom. Ti primeri navezovanja niso nikakor znamenje posebnega fonološkega ali morfološkega statusa »veznega soglasnika«. Vseeno pa je zanimivo, da v primeru nekaterih pridevniških priredij vloge fonologije pri izbiri OVMS ne gre zanemariti: te oblike so sorazmerno manj številne pred soglasnikom kot pred samoglasnikom, a so številnejše, ko se z njimi lahko ognemo dvojnemu hiatu.



## THE IMPERATIVE IN NATURAL SYNTAX

Natural Syntax is a pseudo-deductive linguistic theory, and this is its most recent version. Natural Syntax determines the presuppositions on the background of which a (morpho)syntactic state of affairs can be made predictable, and thus synchronically explained. The two basic kinds of presuppositions are what are known as naturalness scales and rules of alignment among corresponding values of any two scales. Every (morpho)syntactic state of affairs is represented by two comparable variants. Natural Syntax contains no generative component.

Natural Syntax is a special case of the approach usually called Naturalness, which arose as a reaction against the pervasive abstractness of generative grammars. (This abstractness was first of all embodied in the distinction between deep and corresponding surface structures). Initially Naturalness was active in phonology (Stampe 1979) and – to a much greater extent – in morphology (Mayerthaler 1981 and amply in Dressler; e.g., Kilani-Schoch/Dressler 2005). The present author added syntax about 1985. Natural Syntax (as my framework is dubbed) introduced a new set of basic criteria and confined the processed language material to what could be assumed to constitute pairs of variants (rather than single structures). Natural Syntax is a deductive theory except that the basic criteria are not couched in mathematical terms, and thus are not true axioms. Therefore Natural Syntax is more correctly described as pseudo-deductive.

The present state of the art can be summarized as follows. Natural Syntax does not perform as efficiently as generative grammars, which circumstance is partly counterbalanced by the much simpler apparatus of Natural Syntax. Natural Syntax is nonetheless interesting in its own right because it has developed a few (not language-particular) distinctions presumably lacking in generative grammars. Restricting myself to what is presented in greater detail below, I mention natural and unnatural environments (crucially combined with parallel and chiasmic alignments).

The basic format of my naturalness scales is  $\text{>nat}(A, B)$ , in which  $A$  is more natural than  $B$ . Two expanded scales are allowed, viz.  $\text{>nat}(A + B, B)$  and  $\text{>nat}(A, A + B)$ ; they are valid if the corresponding scale of the format  $\text{>nat}(A, B)$  is valid. Exemplification below.

I proceed to list the criteria with which Natural Syntax substantiates naturalness scales. (The following basic criteria (a–h) are sometimes referred to as axioms in this paper in order to draw attention to the circumstance that the criteria are presupposed.)

---

\* *Author's address:* Filozofska fakulteta, Oddelek za primerjalno in splošno jezikoslovje, Aškerčeva 2, Ljubljana, Slovenia. Email: janez.oresnik@sazu.si

- (a) The speaker/hearer parameter. In the scale >nat (A, B), value A is natural for the speaker (and unnatural for the hearer); value B is unnatural for the speaker (and natural for the hearer). The basic naturalness scale is >nat (favourable for the speaker, favourable for the hearer). This view of naturalness is commonplace in linguistics (Havers 1931: 171), under the names of tendency to economize (utilized first of all by the speaker) and tendency to be accurate (mainly in the hearer's interest).

I follow Mayerthaler (1981: 13 ff.) in assuming that the speaker is the centre of communication, and therefore most properties of the speaker are natural; for instance, being the first person and/or the subject and/or +human and/or +masculine (!) and/or +singular and/or +definite and/or +referential, etc.

What is favourable for the hearer may be less natural for the speaker. This is a pivotal point in Natural Syntax and will be maintained until some good counterexample nullifies it. By way of illustration it can be pointed out that producing a longish noun phrase may be "tiresome" for the speaker (= less natural for him), but may ease the hearer's decoding process considerably (= be more natural for the hearer).

- (b) The principle of least effort (Havers 1931: 171). What conforms better to this principle is more natural for the speaker. What is cognitively simple (for the speaker) is easy to produce, easy to retrieve from memory, and so on.
- (c) Degree of integration into the construction. What is better integrated into its construction is more natural for the speaker.

As a rule of thumb, what is located at the margin of a construction is less natural (for the speaker) than what is placed inside a construction.

- (d) Frequency. What is more frequent tokenwise is more natural for the speaker. What is cognitively simpler (for the speaker) is used more. (However, the reverse does not obtain: what is natural for the speaker is not necessarily more frequent.)
- (e) Small vs. large class. The use of (a unit pertaining to) a small class is more natural for the speaker than the use of (a unit pertaining to) a large class. During speech small classes are easier for the speaker to choose from than are large classes. (This is frequency typewise.)
- (f) The process criterion. Any process is natural. Examples of processes: movement, agreement.
- (g) Acceptable vs. non-acceptable use. What is acceptable is more natural for the speaker than what is not acceptable. The very reason for the acceptability of a syntactic unit is its greater naturalness for the speaker with respect to any corresponding non-acceptable unit.
- (h) What is more widespread in the languages of the world is more natural for the speaker (the typological criterion). What is cognitively simpler (for the speaker) is realized in more languages.

I have been applying the above criteria (a–h) to language material covering several languages and miscellaneous (morpho-)syntactic states of affairs. Throughout my work, the criteria have compelled me, time and again, to reject certain solutions and to give precedence to other solutions. Given this encouraging experience, I will preserve the present list (a–h) until some convincing and irreparable counterexample casts doubt upon my axioms. The occurrence of such an event is in the overriding interest of Natural Syntax anyway. The only realistic aim of deductive theories is that they are eventually disproved. I am afraid that any improvement of the axioms would lead to a reduction of the chances for the desirable definitive outcome.

The above criteria of naturalness are utilized to support my naturalness scales. Normally it suffices to substantiate any scale with one criterion, which backs up either value A or value B of the scale; the non-supported value is allotted the only remaining position in the scale. Of course, a scale may be supported with more than one criterion. Any clash among the criteria applied to a scale is to be handled with constraints on the combinations of criteria. So far no convincing constraints have been formulated; I have not yet encountered much useable crucial language data.

The naturalness scales are an essential part of what are known as deductions, in which Natural Syntax expresses its predictions about the state of affairs in language data. An example of a deduction:

English. The numerical indication of frequency normally consists of a cardinal number followed by the word *times* (e.g., *four times*) except that there are one-word expressions available for the lowest numbers: *once*, *twice*, and archaic *thrice* (*Collins Cobuild* 1990: 270–271).

The two variants: the type *once* and the type *four times*.

1. The assumptions of Natural Syntax:

1.1. >nat (type *once*, type *four times*)

I.e., the type *once* is more natural than the type *four times*. – According to the criterion of least effort, item (b) in the list of axioms.

1.2. >nat (low, non-low) / number

I.e., any low number is more natural than any non-low number (Mayerthaler 1981: 15). – Low numbers are more easily accessible to the speaker. According to the speaker/hearer criterion, item (a) in the list of axioms.

2. The rules of parallel alignment of corresponding values:

2.1. value A tends to associate with value C,

2.2. value B tends to associate with value D. See Note 4.1 below.

3. The consequences:

If the language distinguishes between low and non-low numbers in numerical indications of frequency such that one kind of number uses the pattern *four times* and the other kind of number uses the pattern *once*, it is the low numbers that tend to use the pattern *once* and it is the non-low numbers that tend to use the pattern *four times*. Q.E.D. (The reverse situation is not expected.)

#### 4. Notes

- 4.1. Value A of scale 1.1 (= the type *once*) tends to combine with value C of scale 1.2 (= low number). Value B of scale 1.1 (= the type *four times*) tends to combine with value D of scale 1.2 (= non-low number). Similarly in the remaining deductions, with the proviso that the alignment (unlike here) can be chiasitic. Chiasitic alignment is explained below.
- 4.2. Natural Syntax cannot predict the cut-off point between low and non-low numerals.
- 4.3. Henning Andersen (p.c.) has pointed out to me that there is a parallel system covering numerical indications of frequency, *one additional time, two/three/four additional times*, etc., which does not make use of the dichotomy treated in the above deduction. Donald Reindl (p.c.) has added *one more time, two/three/four more times*, etc.
- 4.4. In item 3 the use of the verb *tend* is important. The principled implication is that any language phenomenon can have exceptions.

This deduction maintains that the state of affairs cannot be the reverse; i.e., that the numerals above *two* (or *three*) would be one-word formations and that the numerals under *three* (or *four*) would be two-word formations. All predictions of Natural Syntax are restricted to such modest claims about the unlikelihood of the reverse situation.

The following are the deduction-internal presuppositions that must be accepted for the consequences (in fact, predictions) stated in item 3 of any deduction to obtain:

- (i) The description of the language data adopted in the deduction;
- (ii) The choice of the two variants treated in the deduction;
- (iii) The choice of the natural or the unnatural environment in which the deduction proceeds;
- (iv) The choice of the naturalness scales that form the basis of the computation within the deduction.

In every deduction, the rules of alignment play a prominent role; compare item 2 in the above deduction. The alignment rules regulate the combinations of corresponding values of the two naturalness scales mentioned in the deduction.

The alignment can be parallel or chiasitic. Suppose that the two scales are  $\text{>nat}$  (A, B) and  $\text{>nat}$  (C, D). Parallel alignment pairs value A with value C, and value B with value D. Chiasitic alignment pairs A with D, and B with C.

A paramount question is when the alignment is parallel and when chiasitic. Parallel alignment is the default case. Chiasitic alignment is necessary whenever a given deduction is limited to the language data obtaining within an "unnatural environment". This is defined as value B of the scale  $\text{>nat}$  (A, B).

An example: in the scale  $\text{>nat}$  (main, dependent) / clause, the value "dependent clause" is an unnatural environment. This means: all deductions whose language data lie within the environment "dependent clause" require the implementation of chiasitic alignment.



Chiastic alignment is prohibited when a naturalness scale is substantiated with an axiom. If, however, an axiom is engaged as one of the scales in a deduction, it obeys the usual distribution of the alignment rules. Although Natural Syntax in principle does not deal with semantic phenomena, it does happen sometimes that semantics is involved in a deduction. My experience so far suggests that semantic phenomena block chiastic alignment within such a deduction. See section (C) below.

The insistence of Natural Syntax on the distinction between parallel and chiastic alignments stems indirectly from the work of Henning Andersen within markedness theory. Andersen observes situations such as the following in all human semiotic systems: on an everyday occasion casual wear is unmarked, and formal wear marked; on a festive occasion it is the formal wear that is unmarked, whereas casual wear is marked. See Andersen (1972: 45, esp. fn. 23). This example expressed with Natural Syntax scales: (i) >nat (casual, formal) / wear, (ii) >nat (-, +) / marked. A third scale as the source of the environment of the deduction: >nat (everyday, festive) / occasion. If the environment is "everyday occasion", the alignment within (i–ii) is parallel; if the environment is "festive occasion", the alignment within (i–ii) is chiastic.

The purpose of this paper is twofold: (i) to acquaint the reader with the application of Natural Syntax a step outside its original domain; namely, in morphology; and (ii) to demonstrate how Natural Syntax treats the morphology and in part the syntax of the imperative. To facilitate the reader's initiation, the illustration is as simple as possible. A brief conclusion ends the paper.

The imperative is almost ubiquitous and its functions display relatively little variation. Yet the imperative is a virtually unsolved linguistic riddle regarding where it belongs. Some linguists advocate the complete independence of the imperative, some pair it with the indicative, and some with the non-indicative moods. Natural Syntax cannot be of help in this quandary; it can only determine whether the imperative is a less or more natural category. From the answer to this question it could follow whether the imperative is less or more natural than the indicative mood, and so on.

I shall proceed in the format of deductions:

(1) Slovenian. There are a synthetic and several analytic imperatives; for instance, *delaj* 'work (2sg)' and *daj delati* give to-work 'work (2sg)'. The analytic imperatives are used only in the colloquial language (Toporišič 2000: 519).

The two variants: the synthetic and the analytic imperative. – The deduction proceeds in the unnatural environment "imperative".

1. The assumptions of Natural Syntax:

1.1. >nat (synthetic, analytic) / imperative

I.e., the synthetic imperative is more natural than the analytic imperative. –

According to the criterion of least effort, item (b) in the list of axioms.

1.2. >nat (+, -) / colloquial language

I.e., the colloquial language is more natural than the non-colloquial language. – Many languages do not use non-colloquial language or use it sparingly. According to the typological criterion, item (h) in the list of axioms.  
A special case of 1.2:

1.2.1. >nat (+, +/-) / colloquial language

Scale 1.2.1 has the permitted format >(A, A + B) and is valid because the corresponding basic scale 1.2 has been substantiated.

2. The rules of chiastic alignment:

2.1. value A tends to associate with value D,

2.2. value B tends to associate with value C.

3. The consequences:

If the language distinguishes between the synthetic and the analytic imperative such that one option is used in the colloquial language and the other option is used both in the colloquial and in the non-colloquial language, then it is the analytic imperative that tends to be used in the colloquial language, and it is the synthetic imperative that tends to be used both in the colloquial and the non-colloquial language. Q.E.D. (The reverse situation is not expected.)

In this deduction, only the imperatives and the (non-)colloquial language cover the entire deduction. The colloquial language is natural, the non-colloquial language is unnatural; thus the default value is natural. Under this assumption, the only trigger of chiastic alignment in deduction (1) can be the synthetic and the analytic imperatives. Consequently the synthetic imperative (like the analytic imperative) must be an unnatural category. This is somewhat surprising, given the known shortness of the synthetic imperative's sound body, at least in the second person singular. However, additional deductions will corroborate this newly obtained conclusion.

(2) Italian. The second person singular imperative. The affirmative imperative ends in a special desinence whereas the negative imperative is formed on the infinitive; for instance, *canta* 'sing (sg.)', *non cantare* 'don't sing (sg.)' (Maiden/Robustelli 2000: 247).

The two variants: the affirmative and the negative imperative of the second person singular. – The deduction proceeds in the unnatural environment "imperative".

1. The assumptions of Natural Syntax:

1.1. >nat (affirmative, negative) / imperative

I.e., the affirmative is more natural than the negative. – The affirmative is usually zero coded in the languages of the world, and therefore the affirmative is natural according to the criterion of least effort, item (b) in the list of axioms.

1.2. >nat (+, -) / infinitive = imperative

I.e., an infinitive expressing an imperative is more natural than a normal infinitive. – Decoding the imperative hidden in an infinitive is difficult for the hearer, and therefore

infinitive = imperative must be mentioned in slot A of the scale. According to the speaker/hearer criterion, item (a) in the list of axioms.

2. The rules of chiastic alignment:

2.1. value A tends to associate with value D,

2.2. value B tends to associate with value C.

3. The consequences:

If the language distinguishes (in the second person singular) between the affirmative and the negative imperatives such that one option assumes the form of an infinitive and the other option does not assume that form, then it is the affirmative imperative that tends not to assume the form of an infinitive and it is the negative imperative that tends to assume the form of an infinitive. Q.E.D. (The reverse situation is not expected.)

4. Note. Excluded from the deduction is the generic infinitive = imperative, acceptable as it is even in the affirmative form: *spingere* 'push', *tirare* 'pull' (signs on doors; Maiden/Robustelli 2000: 248, and Tjaša Miklič p.c.).

(3) Swahili. The imperative ends in *a*; for instance, *soma* 'read (sg.)'. If the imperative has an affix attached, the form ends in *e*; for instance, *some-ni* 'read (pl.)', *vi-some* 'read (sg.) them' (Perrott 1951: 46).

The two variants: the imperatives *soma* and *some*. – The deduction proceeds in the unnatural environment "imperative".

1. The assumptions of Natural Syntax:

1.1. >nat (*some*, *soma*) / imperative

I.e., the form *some* is more natural than the form *soma*. – The form *some* is better integrated into its full form (because of its obligatory affixes), and therefore natural according to the criterion of integration into construction, item (c) in the list of axioms.

1.2. >nat (-, +) / affix

I.e., the absence of an affix is more natural than its presence. – According to the criterion of least effort, item (b) in the list of axioms.

2. The rules of chiastic alignment:

2.1. value A tends to associate with value D,

2.2. value B tends to associate with value C.

3. The consequences:

If the language distinguishes between the imperative variants *soma* and *some* such that one variant has affixes attached and the other variant lacks any affixes, then it is *some* that tends to have affixes attached and it is *soma* that tends to lack any affixes. Q.E.D. (The reverse situation is not expected.)

(4) Turkish. The imperative ending in *a* (disobeys a sound-law and) does not replace the *a* with *ı* before desinence-initial *y*. For instance, *başla-yın* 'begin (pl.)' does not change to *başlı-yın* (Lewis 1953: 80).

The two variants: the imperative forms *bašlî-yîn* in *bašla-yîn*. – The deduction proceeds in the unnatural environment "imperative".

1. The assumptions of Natural Syntax:

1.1.  $\text{>nat} (bašlî-yîn, bašla-yîn) / \text{imperative forms}$

I.e., the form *bašlî-yîn* is more natural than the form *bašla-yîn*. – The form *bašlî-yîn* exhibits repetition whereas the form *bašla-yîn* lacks repetition. Repetition is a kind of imitation; humans have inherited imitation from higher mammals, and therefore imitation is very natural.

1.2.  $\text{>nat} (+, -) / \text{acceptable}$

I.e., acceptable is more natural than unacceptable. – This is the very criterion of acceptability, item (g) in the list of axioms.

2. The rules of chiasitic alignment:

2.1. value A tends to associate with value D,

2.2. value B tends to associate with value C.

3. The consequences:

If the language distinguishes between the imperative variants *bašlî-yîn* and *bašla-yîn* such that one variant is acceptable and the other variant is not acceptable, then it is the variant *bašlî-yîn* that tends not to be acceptable and it is the variant *bašla-yîn* that tends to be acceptable. Q.E.D. (The reverse situation is not expected.)

The assumption that the imperative is an unnatural category clashes (at least apparently) with the circumstance that the following combination of the Slovenian imperative with the grammatical number requires parallel alignment:  $\text{>nat} (teci, tecita/-te)$  'run (sg.) [vs] run (du.), run (pl.)' and  $\text{>nat} (\text{singular, non-singular})$ . Chiasitic alignment is expected, seeing that the deduction just set lies wholly within the unnatural environment "imperative", yet in reality, I repeat, only parallel alignment yields the correct result. To justify parallel alignment it must be stipulated that the former scale is arbitrarily restricted to the imperative. This is supported by the following deliberation:

Given the scale  $\text{>nat} (\text{singular, non-singular})$ , adduced above, the non-singular is an unnatural environment, and as such requires chiasitic alignment. Consider now the agreement (of the subject with the finite verb) within the indicative of the present tense. (I continue to use Slovenian as example.) The following pairs of scales are relevant:

(5)  $\text{>nat} (\text{singular, dual}) / \text{subject}$ ,  $\text{>nat} (\text{singular, dual}) / \text{finite verb}$ . This means: if the subject is in the singular or the dual, even the corresponding finite verb is in the singular or the dual, respectively. The alignment is parallel, as expected in a deduction not restricted to an unnatural environment. The two scales correctly predict the situation in the language material.

(6) >nat (singular, plural) / subject, >nat (singular, plural) / finite verb. Everything is comparable to item (5), including the alignment (it is parallel). Again, the two scales correctly predict the situation in the language material.

(7) >nat (plural, dual) / subject, >nat (plural, dual) / finite verb. This deduction proceeds in the unnatural environment "non-singular" (mentioned above) and chiasitic alignment is required. More precisely, the plural subject agrees with the finite verb in the dual; the dual subject agrees with the finite verb in the plural. As can be seen, the deduction predicts a false situation in the language material. For this reason the scales under (7) must be removed from Natural Syntax, and this operation should be undertaken in a principled way.

I see a solution in the experiment to encompass the agreement of all subjects and all finite verbs in one and the same deduction:

(8) >nat (singular,  $\alpha$ non-singular) / subject, >nat (singular,  $\alpha$ non-singular) / finite verb. In both scales,  $\alpha$  stands for the dual or plural. The alignment must be parallel seeing that the deduction is not restricted to any unnatural environment. The prediction is correct: if the subject is in the singular, the finite verb is likewise in the singular, and so on. The deduction takes care of the imperative forms as well and, again, correctly predicts the situation in the language material.

The principled character of this move is reflected in the circumstance that the new solution leaves behind the problem connected with the imperative and with deductions (5–7) in the sense that the problem becomes a special case of the solution not to be separately mentioned in the solution. This is achieved by stipulating that a good solution tends to encompass as much relevant language material as possible, depending, of course, on the potential of the formalism of Natural Syntax.

There follows another example involving the imperative and the grammatical number, yet completely restricted to the imperative, and consequently quite different from the previous matter:

(9) Old French. The singular forms of the imperative are different from the corresponding indicative forms (*chante* 'sing (sg.)', *dorm* 'sleep (sg.)', *vent* 'sell (sg.)'), whereas the plural forms of the imperative equal the corresponding indicative forms (Bourciez 1967: 332).

The two variants: the singular and the plural of the imperative. – The deduction proceeds in the unnatural environment "imperative".

1. The assumptions of Natural Syntax:

1.1. >nat (singular, plural) / imperative

I.e., the singular is more natural than the plural. – In many languages the singular is zero coded and therefore natural according to the criterion of least effort, item (b) in the list of axioms.

1.2. >nat (+, -) / imperative = indicative mood

I.e., any imperative in the form of the indicative is more natural than any imperative different from the indicative mood. – The decoding of the imperative from the indicative mood is difficult for the hearer, and therefore imperative = indicative must be mentioned in slot A of the scale. According to the speaker/hearer criterion, item (a) in the list of axioms.

2. The rules of chiasitic alignment:

2.1. value A tends to associate with value D,

2.2. value B tends to associate with value C.

3. The consequences:

If the language distinguishes between the imperative's singular and plural such that one grammatical number equals the indicative and the other grammatical number does not equal the indicative, then it is the singular that tends not to equal the indicative and it is the plural that tends to equal the indicative. Q.E.D. (The reverse situation is not expected.)

I proceed to the question of why the imperative (at least but often not only of the second person singular) is conspicuously short in most languages. There even used to be a fashion in structuralist morphology to derive the forms of any verb from its imperative as a basis, if only the imperative lacked any desinences. Natural Syntax cannot explain how the shortness of the imperative comes into being; this issue pertains to historical grammar and to the generative component (if available) of descriptive grammar. However, Natural Syntax has a legitimate interest in why the imperative has remained persistently short for centuries if not millennia, and quite frequently manages to avoid the lengthening of its sound body. The immediate answer is that the shortness of the imperative must be descriptively motivated (must have a deeper sense, descriptively speaking). I propose to approach a detailed answer through a discussion about zero case endings of the Slovenian noun. If the noun has a zero ending in the nominative singular (e.g., *klobuk-∅* 'hat'), that ending does not cause much surprise because the nominative is an unmarked case; this circumstance makes the *-∅* sufficiently motivated. There is a different situation in the genitive non-singular when it displays a zero ending; for instance, *žab* from *žaba* 'frog', *konj* from *konj* 'horse', *mest* from *mesto* 'town', and *volov* from *volovi* 'oxen'. What is the motivation of such a short genitive? The language is certainly able to change a genitive of this kind; for instance, in addition to the genitive *konj-∅* 'of horses' there is now also *konj-ev*. However, the language does not always act thus, but instead preserves the short form endowing it – I suggest – with a special motivation: any unexpectedly short form is classified as an ordinary (thus stable) word after it has received a special feature; let the feature be called a "shortened word". Shortened words are stable words that are permitted to live a long life in spite of their shortness. Natural Syntax assigns the fea-

ture "shortened word" by aid of the scale  $\gt$ nat (shortened, ordinary) / word. Consider the deduction in which the genitive *žab* is assigned that feature:

(10) Slovenian. The noun *žaba* 'frog' has a case ending in all its forms except in the genitive non-singular: *žab*.

The two variants: *žab* + case ending and the bare *žab*.

1. The assumptions of Natural Syntax:

1.1.  $\gt$ nat (bare *žab*, *žab* + case ending)

I.e., the bare *žab* is more natural than *žab* + case ending. – According to the criterion of least effort, item (b) in the list of axioms.

1.2.  $\gt$ nat (shortened, ordinary) / word

I.e., a shortened word is more natural than an ordinary word. – According to the criterion of least effort, item (b) in the list of axioms.

2. The rules of parallel alignment:

2.1. value A tends to associate with value C,

2.2. value B tends to associate with value D.

3. The consequences:

If the language distinguishes, within the inflexion of the noun *žaba*, between the bare *žab* and *žab* + case ending such that one option is a shortened word and the other option is an ordinary word, then it is the bare *žab* that tends to be a shortened word and it is the form *žab* + case ending that tends to be an ordinary word. Q.E.D. (The reverse situation is not expected.)

4. Notes

4.1. It can be seen in scale 1.1 that the feature "shortened form" can be assigned whenever such a form undergoes pressure from the remaining forms of the paradigm.

4.2. Two further examples. The Latvian vocative *siev* 'madam' beside the longer nominative *sieva* 'mistress' (Mayrhofer 1986: 149). The Indo-European vocative of  $\bar{a}$ -stems was a shortened form in  $-\bar{a}$  (subsequently  $-o$  in Slavic languages; for instance, Croatian vocative *žen-o* from *žena* 'woman').

4.3. In Serbian and Croatian Štokavian dialects the inherited genitive *žab* did not obtain the status of a shortened word nor was it given any other support, and the form underwent renewal so as to become an ordinary word: *žab-ā*.

The genitive plural *žab* is an important example. The zero case ending resulted from regular sound changes in the history of the language and has been preserved throughout the last millennium. The form does not display any tendency to develop a non-zero ending. This circumstance happens to remind the author of the two Slovenian verbs *nesti* and *nositi*, both meaning 'to carry', whose root vowels *e* and *o* are inherited from Proto-Indo-European. The original motivation for the vowel difference ceased to be active thousands of years ago, yet there is no inclination towards change: for instance, towards the neutralization of the vowel difference. In both cases (i.e., with the genitive *žab* and with the pair *nesti/nositi*) the descriptive grammar must explain

why the two cases are stable or, more precisely, what deeper sense they are ascribed in the brains of native speakers of Slovenian. The pair *nesti/nositi* will not be discussed here; the genitive *žab* has just been interpreted in the framework of Natural Syntax.

After this discussion on frogs, I return to the imperative. To shortened words will now be added shortened constructions.

(11) English. A covert subject of the imperative corresponds to the unaccented personal-pronoun subject of non-imperative finite verbs: *you go* (*you* unaccented) is not acceptable.

The two variants: imperative *go* and non-imperative *you + go*.

1. The assumptions of Natural Syntax:

1.1. >nat (imperative *go*, non-imperative *you + go*)

I.e., *go* is more natural than *you + go*. – According to the criterion of least effort, item (b) in the list of axioms.

1.2. >nat (shortened, ordinary) / word/construction

I.e., a shortened word/construction is more natural than an ordinary word/construction. – According to the criterion of least effort, item (b) in the list of axioms.

2. The rules of parallel alignment:

2.1. value A tends to associate with value C,

2.2. value B tends to associate with value D.

3. The consequences:

If the language distinguishes between the imperative *go* and the non-imperative *you + go* such that one option is a shortened word/construction and the other option is an ordinary word/construction, then it is the imperative *go* that tends to be a shortened word/construction and it is the non-imperative *you + go* that tends to be an ordinary word/construction. Q.E.D. (The reverse situation is not expected.)

4. Notes

4.1. It can be seen in scale 1.1 that the feature "shortened construction" can be assigned whenever such a construction undergoes pressure from the remaining constructions of the paradigm.

4.2. Accordingly the imperative *go* is a shortened word/construction. What has been removed is its pronominal subject = the unaccented *you*.

4.3. This deduction cannot predict the exact shape of any single shortened word/construction.

4.4. A different deduction could predict that the imperative *you go* (*you* unaccented) is not acceptable.

4.5. Deduction (11) is valid for all languages featuring obligatory pronominal subjects; thus in Europe for the Germanic and in part the Romance languages. For the Icelandic case cf. Note 4.6.



- 4.6. An Icelandic example: the old two-word imperative *ger þú* (*þú* unaccented [!]) 'do (sg.)' was shortened to the present single-word *gerðu*; similarly with all imperatives of the second person.

(12) German. The infinitive can be used as a generic imperative: *aufpassen* 'take care'. With the reflexive verbs (for example, *sich beeilen* 'hurry up', *sich hinauslehnen* 'lean out') the imperative use of the infinitive lacks the reflexive pronoun; for instance, *bitte beeilen* 'please hurry up', *nicht hinauslehnen* 'don't lean out', warning under window of train (*Duden. Richtiges und gutes Deutsch* 1985: 361).

The two variants: the infinitive *sich beeilen* and the imperative *beeilen*.

1. The assumptions of Natural Syntax:

1.1. >nat (+, -) / infinitive = imperative

I.e., an infinitive expressing an imperative is more natural than a normal infinitive. – Decoding the imperative hidden in an infinitive is difficult for the hearer, and therefore infinitive = imperative must be mentioned in slot A of the scale. According to the speaker/hearer criterion, item (a) in the list of axioms.

1.2. >nat (shortened, ordinary) / word/construction

I.e., a shortened word/construction is more natural than an ordinary word/construction. – According to the criterion of least effort, item (b) in the list of axioms.

2. The rules of parallel alignment:

2.1. value A tends to associate with value C,

2.2. value B tends to associate with value D.

3. The consequences:

If the language distinguishes (with reflexive verbs) between an imperative and a non-imperative infinitive such that one option represents a shortened word/construction and the other option represents an ordinary word/construction, then it is the imperative infinitive that tends to represent a shortened word/construction and it is the non-imperative infinitive that tends to represent an ordinary word/construction. Q.E.D. (The reverse situation is not expected.)

4. Notes

4.1. The imperative infinitive, being a shortened word/construction, has lost the reflexive pronoun.

4.2. The reflexive pronoun is removed from the imperative infinitive by aid of the scale >nat (-, +) / reflexive pronoun, to replace item 1.2.

Imperative clauses can be long or short; for instance, *go (where destiny calls you)*. Yet short imperative clauses are in the majority and it may be right to assign them the status of shortened constructions.

Shortened constructions are also to be found outside imperative clauses. Suffice it to mention the prevailing short reporting clause of reported speech; for instance, *John said that it was too late*.

It happens less often that an second person singular imperative is lengthened:

(13) Modern Greek. The second person singular of the imperative 'say' ought to be *pe*, but the true form is *pe-s*. (Thus only in a few verbs; Sofroniou 1962: 91.)

The two variants: the imperatives *pe* in *pes*. – The deduction proceeds in the unnatural environment "imperative".

1. The assumptions of Natural Syntax:

1.1. >nat (*pe*, *pes*) / imperative

I.e., the form *pe* is more natural than the form *pes*. – According to the criterion of least effort, item (b) in the list of axioms.

1.2. >nat (+, -) / acceptable

I.e., acceptable is more natural than unacceptable. – This is the very criterion of acceptability, item (g) in the list of axioms.

2. The rules of chiasitic alignment:

2.1. value A tends to associate with value D,

2.2. value B tends to associate with value C.

3. The consequences:

If the language distinguishes between the imperative variants *pe* and *pes* such that one variant is acceptable and the other variant is not acceptable, then it is the variant *pes* that tends to be acceptable and it is the variant *pe* that tends not to be acceptable. Q.E.D. (The reverse situation is not expected.)

4. Notes

4.1. The extension *-s* was in use with a few (aorist) imperatives as early as Ancient Greek; for instance, *skhē-s* 'get, obtain (sg.)'. The origin of the extension is not clear (Sihler 1995: 602).

4.2. The second person singular imperative of Proto-Indo-European athematic (thus shortest possible) verbs probably originally consisted of the bare root frequently accompanied by the particle *dhi*. The imperative + particle *dhi* was in time shortened to the imperative ending in *-dhi*. (With thematic verbs the imperative of the second person singular was likewise a shortened word equal as it was to the verbal stem.)

4.3. In Lithuanian all imperatives are extended with *-k(-)*; for instance, *sup-k* 'rock (2nd sg.)', *sup-ki-me* 'let us rock', *sup-ki-te* 'rock (2nd pl.)' (Ambrazas ed. 1997: 315). This phenomenon is not limited to the second person singular and only resembles the modern Greek state of affairs. Notice that the extension *-k(-)* is not followed by anything else in the second person singular, which is thus a shortened word.

The notion "shortened word/construction" (the expression has an ad hoc character) is meant to be a grammatical category. I propose that the status of a word/construction as a shortened word/construction is evaluated by the brain as relevant information and is forwarded to subsequent generations of native speakers of that language during the acquisition of their native language. This guarantees the stability of any shortened word/construction.

## CONCLUSION

The basic idea has been to illustrate how a (pseudo-)deductive theory of syntax performs if it insists on avoiding abstract solutions, and in particular on excluding any generative component. The main limitation of such an approach seems to consist in the compulsory reliance on language material that contains pairs of variants (aspects of whose syntactic behaviour can be made predictable using suitable presuppositions). Single constructions have no place in Natural Syntax.

In this paper Natural Syntax has been applied to the morphology and in part the syntax of the imperative. Apparently the framework can be successful even in an area outside syntax proper.

## References

- AMBRAZAS, Vytautas (ed.) (1997) *Lithuanian grammar*. Vilnius: Baltos lankos.
- ANDERSEN, Henning (1972) "Diphthongization." *Language* 48, 11–50.
- BOURCIEZ, Édouard (1967) *Éléments de linguistique romane*. Fifth edition. Paris: Klincksieck.
- Collins Cobuild English Grammar*. London: HarperCollins, 1990.
- CVETKO-OREŠNIK, Varja/Janez OREŠNIK (2009) "Natural Syntax: Expressions of future time." In: Franks et al. (eds.): *A linguist's linguist. Studies in South Slavic linguistics in honor of E. Wayles Browne*. Bloomington: Slavica, 167–87.
- Duden. Richtiges und gutes Deutsch*. Third edition. Mannheim: Dudenverlag, 1985.
- HAVERS, Wilhelm (1931) *Handbuch der erklärenden Syntax*. Heidelberg: Winter.
- KILANI-SCHOCH, Marianne/Wolfgang U. DRESSLER (2005) *Morphologie naturelle et flexion du verbe français*. Tübingen: Narr.
- LEWIS, Geoffrey L. (1953) *Teach yourself Turkish*. London: The English Universities Press.
- MAIDEN, Martin/Cecilia ROBUSTELLI (2000) *A reference grammar of modern Italian*. London: Arnold.
- MAYRHOFER, Manfred (1986) *Indogermanische Grammatik*. Heidelberg: Winter.
- OREŠNIK, Janez (2009a) "Transitivity in Natural Syntax: Accusative languages." *Poznań Studies in Contemporary Linguistics* 45, 405–446.
- OREŠNIK, Janez (2009b) "Transitivity in Natural Syntax: Ergative languages." *Linguistica* 49, 65–93.
- OREŠNIK, Janez (2009c) "Natural Syntax of Belfast English (I) Subject-verb agreement (II) Imperative." *Studia Anglica Posnaniensia* 45(2), 107–143.
- PERROTT, Daisy Valerie (1951) *Teach yourself Swahili*. London: The English Universities Press.
- SIHLER, Andrew L. (1995) *New comparative grammar of Greek and Latin*. New York: Oxford University Press.
- SOFRONIOU, Sofronios Agathocli (1962) *Modern Greek*. London: Teach Yourself Books.
- STAMPE, David (1979) *A dissertation on Natural Phonology*. New York: Garland.
- TOPORIŠIČ, Jože (2000) *Slovenska slovnica*. Fourth edition. Maribor: Obzorja.

Abstract  
THE IMPERATIVE IN NATURAL SYNTAX

The framework of this paper is Natural Syntax initiated by the author in the tradition of (morphological) naturalness as established by Wolfgang U. Dressler and †Willi Mayerthaler.

Natural Syntax is a pseudo-deductive linguistic theory, and this is its most recent version. The naturalness judgements are couched in naturalness scales, which follow from the basic parameters (or "axioms") listed at the beginning of the paper. The predictions of the theory are calculated in what are known as deductions, the chief components of each being a pair of naturalness scales and the rules governing the alignment of corresponding naturalness values. Parallel and chiasitic alignment is distinguished and related to Henning Andersen's early work on markedness.

The basic idea is to illustrate how a (pseudo)deductive theory of syntax performs if it insists on avoiding abstract solutions, and in particular on excluding any generative component.

In this paper the framework of Natural Syntax is applied mostly outside its original domain; namely, to the morphological behaviour of the imperative in various languages.

Povzetek  
VELELNİK V NARAVNI SKLADNJI

Naravna skladnja je deduktivna teorija v razvoju in veja teorije naravnosti. Sodbe o naravnosti so ubesedene v lestvicah naravnosti, ki sledijo iz osnovnih meril (ali "aksiomov"), naštetih v začetku sestavka. Napovedi teorije se izračunavajo v t.i. izpeljavah, katerih glavni sestavini sta par lestvic naravnosti in pravila o ujemanju med soodnosnimi vrednostmi naravnosti. Ločimo vzporedno in križno ujemanje, ki sta v dopolnjevalni razvrstitvi. Križno ujemanje je obvezno v izpeljavah, omejenih na nenaravno okolje.

Glavni namen naravne skladnje je določati pogoje, pod katerimi so razmere v nekem jezikovnem gradivu napovedljive.

Sestavek je posvečen velelniku v raznih jezikih, in sicer z oblikoslovnega stališča. (Čista skladnja je upošteta le bolj obrobno.) Velelnik je določen kot nenaravna kategorija. Zlasti oblika za drugo osebo ednine je po jezikih nenavadno kratka in zvečine ne kaže razvojne težnje k podaljšanju. Ta okoliščina je v članku osmišljena tako, da je kratkemu velelniku pripisana (nova) slovnična oznaka "skrajšana beseda", ki se z usvajanjem materinščine prenaša iz rodu v rod naravnih govorcev tistega jezika ter je porok za trajnost (stabilnost) velelnikove kratkosti.

Kot skrajšane besede so v članku prepoznani tudi roditeljski neednine vrste *žab* in *mest*, kot skrajšane zgradbe pa m.dr. spremni stavki poročanega govora.

## LE GÉRONDIF ET LE PARTICIPE PRÉSENT ET LEUR ÉVOLUTION VERS LA GRAMMATICALISATION : ÉTUDE CONTRASTIVE DU SLOVÈNE ET DU FRANÇAIS

### 0. LES PROCÉDÉS DE LEXICALISATION ET GRAMMATICALISATION

Les gérondifs et les participes, formes verbales du mode impersonnel, subissent dans leur manifestation moderne, synchronique, la grammaticalisation aussi bien en français qu'en slovène. Ces formes persistent avec les formes des gérondifs et participes dans leur rôle verbal de base.

La grammaticalisation, un terme qui est (re)devenu à la mode dans les deux dernières décennies, a été utilisée pour la première fois par A. Meillet (1921) dans son ouvrage *Linguistique historique et linguistique générale*. Au chapitre *L'évolution des formes grammaticales*, il mentionne que « l'innovation analogique et l'attribution d'un caractère grammatical à un mot jadis autonome sont les seuls (procédés) par lesquels se constituent des formes grammaticales nouvelles » (ibid. p. 130). Il l'a démontré, entre autres, aussi sur l'exemple du développement de la négation en français à l'aide du mot *pas* (1921 : 139).

La grammaticalisation est le procédé dans lequel une forme lexicale endosse peu à peu le rôle d'une catégorie grammaticale dépendante (Bussman 1990 : 289 – 291) et qui perd sa valeur indépendante. C. Lehmann, qui a traité la grammaticalisation dans ses ouvrages de 1985 et 2005, a passé par une évolution. Il s'est déplacé dans ses recherches vers les développements cognitifs. L'essentiel de ses thèses sur la grammaticalisation est que les catégories majeures comme le nom et le verbe passent vers les catégories secondaires comme l'adverbe ou la préposition. Les critères de la grammaticalisation (selon Lehmann 2005) sont la perte de l'autonomie et la désémantisation. Le procédé inverse à la grammaticalisation est la lexicalisation. Selon Lehmann (1985), il s'agit de « l'entrée dans la grammaire » dans le cadre de la grammaticalisation et de « l'entrée dans le lexique » dans le cadre de la lexicalisation. Pour cette dernière, Bussman (1990 : 452) mentionne, dans la synchronie, l'acceptation du mot dans le dictionnaire (et dans le lexique d'une langue). Dans le cadre de la diachronie, la lexicalisation relève du processus ou du résultat de la démotivation où plusieurs morphèmes sont unis en unité lexicale qui n'est pas toujours déchiffrable des sens composés de ces morphèmes, comme dans l'idiomatization où *passer l'arme à gauche* signifie *mourir*.

---

\* Adresse de l'auteur : Filozofska fakulteta, Oddelek za prevajalstvo, Aškerčeva 2, 1000 Ljubljana, Slovénie. Mél: mojca.brezar@ff.uni-lj.si

Pour C. Marchello-Nizia (2006 : 15),

On nomme « grammaticalisation » un type de changement linguistique très répandu dans toutes les langues du monde. On a coutume, dès le début, de le décrire par son résultat : c'est le processus par lequel les lexèmes deviennent des morphèmes. Ces nouvelles unités grammaticales servent à coder des relations qui n'étaient pas codées grammaticalement auparavant, soit qui l'étaient mais différemment (E. Traugott, E. Koenig 1991: 189).

Selon elle (2006 : 32), la grammaticalisation se situe à trois niveaux :

Au niveau formel, dans le cours du processus de la grammaticalisation, l'évolution se fait soit sans changement de forme soit en allant vers une forme plus réduite.

Au niveau catégoriel, on va toujours d'une catégorie majeure (nom, verbe, adjectif) vers une catégorie mineure.

Au niveau sémantique, le sens lexical évolue vers un sens grammatical plus général et plus abstrait.

Il faut mentionner aussi l'impossibilité d'un retour en arrière d'un tel processus.

La différence entre lexicalisation et grammaticalisation dépend du point de vue adopté. Alors que certains chercheurs considèrent que les deux termes se confondent (voir notamment Halmoy 2003), selon M. Fruyt (2009), tout dépend de quelle manière on entre dans le lexique, à partir de la syntaxe ou de la morphologie. Lors d'une lexicalisation, le lexique s'enrichit d'une unité supplémentaire. Lors d'une grammaticalisation, le mot est désémantisé en une catégorie auxiliaire, relationnelle – la conjonction ou la préposition. La hiérarchie des parties du discours donnera la réponse s'il s'agit d'une grammaticalisation ou d'une lexicalisation.

Les deux procédés entrent en jeu dans le cadre de la catégorie grammaticale du participe présent et du gérondif. Nous esquisserons une étude contrastive du gérondif et du participe français et slovènes qui ont des fonctions semblables, équivalentes dans le cadre des constructions détachées, mais dont l'usage en slovène baisse considérablement plus qu'en français.

Nous nous penchons sur l'hypothèse que les participes et les gérondifs présentent de moins en moins une classe d'activité de la morphologie et, par conséquent, deviennent de plus en plus grammaticalisés. Selon M. Fruyt (2009), les adverbes sont une catégorie plus grammaticalisée que les adjectifs ou les verbes parce qu'ils ne sont pas soumis à la flexion.

D'après C. Marchello-Nizia, il s'agit d'une « dégradation » allant de la forme verbale vers la forme adjectivale et adverbiale (2006 : 36). Nous pouvons en déduire que tout procédé survenu à des gérondifs ou des participes est du côté de la grammaticalisation. Nous tâcherons de définir, à l'aide des dictionnaires et grammaires, jusqu'à quel degré la grammaticalisation a été effectuée en français et en slovène et vérifierons les données à partir des corpus monolingues et bilingues parallèles.

## 1. LE MODE IMPERSONNEL VERBAL ET SA MORPHOLOGIE

Dans cette partie, nous esquisserons la morphologie du gérondif, du participe présent et de l'adjectif verbal en français et en slovène. Quelques-unes de ces formes verbales ont subi des transformations en devenant des adjectifs et des adverbes ou même des prépositions ou des conjonctions, elles ont donc baissé dans la hiérarchie des parties du discours pouvant former le noyau de la prédication pour devenir des mots outils grammaticaux.

### 1.1 Le gérondif en français à travers l'histoire et dans les grammaires contemporaines

En français, le participe et le gérondif sont formés à partir de la base de l'infinitif verbal et de la terminaison *-ant*. Le gérondif présente une nouveauté par rapport au fonds latin.

Selon O. Halmoy (2003 : 37),

l'émergence du gérondif constitue une énigme de l'histoire du français. Dans une perspective diatopique et synchronique on peut voir que le gérondif n'avait d'équivalent morphologique et syntaxique exact ni dans les langues romanes, ni dans les langues germaniques. Il n'en avait pas en latin.

Du latin à l'ancien français, de trois formes latines (*amandum*, *amandi*, *amando*) on a abouti à une forme en *-ant* en ancien français. On rencontre des formes attestées déjà en ancien français (XII<sup>e</sup> – XIII<sup>e</sup> siècle), par exemple dans *La chanson de Roland* et *Le Roman de Tristan (en estant, en seant, en gisant, en riant)* (Halmoy 2003). En moyen français (XIV<sup>e</sup>–XV<sup>e</sup> siècle) il y avait, selon O. Halmoy (2003 : 49) 73 occurrences de gérondifs en 226 pages, ce qui donne une occurrence en moyenne toutes les 3 pages. La situation en français pré-classique (XVI<sup>e</sup> siècle) et le français classique (XVII<sup>e</sup> siècle) est similaire. L'état actuel s'est fixé entre le XIII<sup>e</sup> et XV<sup>e</sup> siècle.

Les grammairiens (p. ex., Riegel et al. 1994 : 339-342) qui considèrent le gérondif comme un syntagme à fonction adverbiale, s'intéressent surtout à la valeur circonstancielle de la tournure qui n'est pas propre au gérondif, mais au lien logique qui s'établit dans le contexte. Ainsi le gérondif peut exprimer une relation de temps – simultanéité (exemple (1), de manière (exemple (2), de cause (exemple (3), de condition (exemple (4), mais aussi de moyen, d'inclusion ou d'équivalence.

- (1) *J'ai pris froid **en attendant** l'autobus.*
- (2) *Ils se sont quittés **en pleurant**.*
- (3) *Je lui ai évité des ennuis **en ne disant rien**.*
- (4) *En **démolissant**, tu lui fais plaisir.*<sup>1</sup>

---

<sup>1</sup> Ces exemples sont inventés.

Le gérondif est apprécié surtout pour sa valeur stylistique. Il apporte de la dynamique dans le récit (Halmoy 2003 : 106) ; il s'emploie pour éviter la lourdeur des phrases, la répétition des relatives, etc. à une contrainte qui doit être respectée, notamment la règle normative de coréférence.

## 1.2 Le participe présent

Ressemblant au gérondif dans ses emplois, le participe est souvent utilisé en remplacement de la construction relative avec *qui*, notamment quand une autre relative est nécessaire dans la phrase (exemple (5)) :

(5) *Nous devons prendre une décision **concernant** l'affaire dont je vous ai parlé la semaine dernière.*

Il est également utilisé pour établir une relation de cause entre deux informations ; ce procédé apparaît surtout dans le style administratif, il sert à mettre en valeur la cause en la présentant comme un argument (exemples (6) et (7)) :

(6) *Ma fille **étant** malade, elle ne pourra assister à votre cours.*

(7) *La direction **ayant refusé** de prendre en considération nos revendications, l'ensemble du personnel se mettra en grève à partir du lundi 8 heures.<sup>2</sup>*

Les participes et gérondifs grammaticalisés sont facilement repérables dans la diachronie. Dans la synchronie, ils présentent la même forme que leurs homologues qui n'ont pas subi le figement. En revanche, ils fonctionnent non pas comme les formes verbales à valeur prédicative mais comme adverbes, conjonctions ou prépositions. O. Halmoy (2003) présente les possibilités suivantes de la grammaticalisation des gérondifs : le figement adverbial, le figement prépositionnel et le figement conjonctionnel avec les exemples suivants : *en attendant*, *en passant* ; *en allant à*, *en passant par*, *en partant de* ; *en admettant (que)*, *en supposant (que)*, *en attendant (que)*.

Nous allons y joindre la recherche de quelques prépositions et conjonctions issues du registre administratif : *étant donné (que)*, *considérant (que)*, *concernant*. Selon C. Marchello-Nizia (2006 : 34), il s'agit de la grammaticalisation, le processus allant d'une partie du discours plus importante vers une partie du discours plus grammaticalisée.

Le participe présent et l'adjectif verbal ont en français dans la plupart des cas une forme identique mais différent en variabilité. La différence en variabilité se fait selon la fonction verbale ou adjectivale<sup>3</sup> et présente un choix arbitraire qui a été normalisé (Wilmet 2007 : 80).

---

<sup>2</sup> Exemples créés pour la démonstration.

<sup>3</sup> Selon Wilmet (2007: 80), l'invariabilité du participe présent était prescrite par un décret du XVII<sup>e</sup> siècle.



Le cas de l'adjectif verbal pouvait être traité comme le figement adjectival total, puisque les adjectifs diffèrent des participes dans l'orthographe (*différent* – *différant* etc.). Des cas similaires sont présents aussi en slovène.

### 1.3 Les gérondifs et les participes présents en slovène

En slovène, on distingue les gérondifs présents en *-č*, *-aje*, *-e* (*videč*, *voyant* – *šepetaje*, *chuchotant* – *molče*, *taisant*) et le participe présent en *-č*. Le participe est variable en genre, en nombre et se décline en fonction du nom avec lequel il entre en relation, tandis que le gérondif reste invariable. Vincenot (1975 : 104) distingue, dans le cadre du gérondif présent en *-č*, la forme adverbiale et la forme adjectivale : en effet, il s'agit de la forme invariable (gérondive) et de la forme du participe présent variable – *-oč* / *-eč*.

Historiquement, les formes gérondives et participiales tirent leur origine du protoslave. En indoeuropéen, le participe présent actif et le gérondif présent étaient formés en *\*-nt*, tandis que le protoslave comprenait trois suffixes : premièrement *\*-y*, *\*-ošti* (la première forme désigne le masculin et le neutre, la deuxième le féminin) ; deuxièmement *-je -jošti* ; et troisièmement *\*-e* et *\*-ęsti* (Jesenšek 1998 : 30 ; Ramovš 1952 : 146-147). Leur usage attesté en slovène a baissé déjà par rapport au slavon et allait dans le sens de la diminution voire l'extinction (Jesenšek 1998 : 39), surtout pour les formes en *-ši* qui ne restent que des adverbes ou des attributs déjà dans la grammaire de Bohorič (1584). Ces formes ont réapparu dans le slovène de l'est au XVIII<sup>e</sup> et XIX<sup>e</sup> siècle. Les gérondifs en *-e* (*molče*, *en se taisant* ; *stoje*, *debout* ; *hote*, *voulant* ; *grede*, *allant*) ont déjà subi l'adverbialisation dès le XVI<sup>e</sup> siècle selon Ramovš (1952). Jesenšek à son tour (1998) situe l'adverbialisation entre les XII<sup>e</sup> et XIII<sup>e</sup> siècles. Le gérondif développe, déjà dans les manuscrits slavons, la faculté de former, dans le cas des sujets identiques, la prédication avec la valeur adverbiale d'une subordonnée circonstancielle (Jesenšek 1998 : 34). Il contribue ainsi à l'économie de l'expression (voir l'exemple (8) ci-dessous) :

- (8) *Pojoč na ves glas so hodili po ulici.*  
*Chantant à haute voix, ils marchaient dans la rue.*

La grammaire du slovène (Toporišič 2000) mentionne à son tour que les gérondifs en *-č* ou *-aje* indiquent l'action verbale qui se déroule en même temps qu'une autre action (simultanément avec une autre action), tandis que le gérondif en *-e* s'emploie aujourd'hui seulement en tant que l'adverbe de manière (exemple (9) ci-dessous) :

- (9) *Molče ves dan je pridno delal.*  
*Silencieusement/ En se taisant, il a travaillé pendant toute la journée.*

Selon Toporišič (1982 : 145, 374, 404), le participe en *-č* dans sa forme adjectivale garde son caractère verbal exprimant la durée de l'action qui se déroule simultanément avec l'activité du sujet et s'accorde en genre, nombre et cas avec le nom – *goreča*

*hiša* (une maison « brûlante » (en feu), *upajoča mati*, une mère « espérante » (dans l'expectative)). Pourtant, certains participes présents ne sont plus que de purs adjectifs. Le figement adjectival est complet, ce qui se voit aussi dans leur sens. Quelques-uns de ces participes d'origine s'éloignent de leur premier sens verbal. Le syntagme *vroča voda* (du participe présent du verbe *vreti* – bouillir) a subi un glissement de sens. Il s'agit aujourd'hui de l'eau *qui est chaude* tandis que le participe signifiait *bouillante*.

Nous pouvons conclure que les gérondifs et les participes sont formés à partir des bases verbales auxquelles s'ajoute le morphème grammatical en fonction de la forme, de l'aspect et de la transitivité du verbe de base. Leur « passage » dans le cadre des mots grammaticaux d'un côté et des nouvelles entrées dans le dictionnaire, de l'autre, se fait par la « dérivation impropre ».

## 2. LE FONCTIONNEMENT DES GÉRONDIFS ET DES PARTICIPES EN FRANÇAIS ET SLOVÈNE CONTEMPORAINS À TRAVERS LES CORPUS

### 2.1 La productivité des formes gérondives et participiales en français et en slovène

En français, le gérondif et le participe restent une ressource stylistique qui évite trop de relativisations et permet d'exprimer un lien temporel, causal ou conditionnel en cas de co-référentialité du sujet. Ces structures sont surtout liées à l'utilisation en français écrit, voire littéraire, soutenu. Le fait que le participe présent et le gérondif ne figurent pas dans les recherches sur l'oral, le prouve. Dans les travaux de M. A. Morel et L. Danon-Boileau (1998) ou C. Blanche-Benveniste (1997), le participe présent et le gérondif sont presque inexistants dans les matériaux étudiés. Pourtant, à l'écrit, cette forme reste bien productive malgré quelques grammaticalisations dans le cadre des gérondifs et participes.

En slovène écrit, le gérondif et le participe *en -č* et *en -e* ont vécu une réintroduction sous influence des écrivains et des journalistes de l'Est de la Slovénie au XVIII<sup>e</sup> et XIX<sup>e</sup> siècle, ce qui a eu pour résultat un allègement syntaxique (Jesenšek 1998 : 40). Toutefois, les formes gérondives et participiales ont été rejetées par les écrivains du début du XX<sup>e</sup> siècle, de sorte que Cankar se tourne vers le slovène parlé cultivé. En même temps, le participe passé en *-ši* est rejeté par les grammaires normatives (Jesenšek 1998 : 53).

Aujourd'hui, le gérondif en *-č*, forme verbale impersonnelle sans réaction à gauche (celle-ci étant déterminée par la co-référentialité du sujet dans la partie verbale et gérondive), formée à partir de la base du présent, marque un événement verbal secondaire se déroulant en même temps que l'événement principal. L'usage décline, étant limité à des textes scientifiques et littéraires surtout en liaison avec des verbes de parole, pensée, perception et mouvement (Jesenšek 1998 : 41).

Une deuxième tentative de la réintroduction apparut dans le domaine de la traduction des textes des institutions européennes (Schlamberger Brezar 2005), mais ce processus a été arrêté avec la standardisation des traductions juridiques et politiques dès l'entrée de la Slovénie en Union européenne.

Dans la suite, nous allons présenter les recherches dans les corpus. Pour le français, nous avons analysé quelques participes et gérondifs choisis où la grammaticalisation a déjà été attestée dans plusieurs emplois tirés des corpus *FraSloK* et *Evrokorpus*, notamment *(en) admettant (que)*, *(en) allant*, *(en) attendant (que)*, *(en) considérant (que)*, *concernant*, qui ont dans la plupart des cas subi la grammaticalisation. Pour le slovène, la recherche a été menée dans le corpus *Fidaplus*<sup>4</sup>.

## 2.2 Les « preuves » de la grammaticalisation des gérondifs et participes

Comme nous l'avons déjà mentionné dans l'introduction, les preuves de la grammaticalisation se font à trois niveaux (C. Marchello-Nizia, 2006 : 32), au niveau formel (la forme se trouvera changée), au niveau catégoriel (du verbe à l'adverbe ou même préposition ou conjonction) et au niveau sémantique, du sens lexical vers un sens grammatical plus général et plus abstrait.

Un autre critère qui peut servir de preuve de la lexicalisation est lié aux propriétés syntaxiques du verbe. Les participes et les gérondifs ont la même propriété que les verbes, qui consiste à fournir le noyau de la construction détachée tandis que les formes adverbialisées ont perdu cette faculté (Toporišič 2000 : 403) : si nous comparons les exemples (10) et (10') où le deuxième énoncé n'est plus normatif en slovène standard :

(10) *Molče je odšel.*

*Il est parti en silence.*

(10') \**Molče ves dan, je delal.*

*En se taisant toute la journée, il a travaillé.*

Un dernier critère peut être le changement d'orthographe et une éventuelle formation des mots nouveaux à partir des gérondifs ou participes lexicalisés. Le participe *rekoč* (*disant*) s'est figé avec sa forme renforcée *tako rekoč* (*soi-disant*) en *tako-rekoč*. Aussi, les participes présents peuvent devenir la base de la formation des adverbes (*razumevajoč* (*qui comprend*) – *razumevajoče*).

La recherche dans le corpus présente une synthèse de tous les éléments cités ci-dessus. Les mots dans leurs contextes et leur fréquence peuvent donner une réponse à la question de la fréquence des participes et gérondifs grammaticalisés et de leur emploi dans différents contextes.

## 2.3. Les participes et gérondifs grammaticalisés à travers les corpus

Puisqu'il n'existe pas de corpus monolingues accessibles pour le français comme c'est le cas pour le slovène avec le corpus de référence *Fidaplus*, nous avons procédé de deux manières différentes pour le slovène et le français. Pour le slovène, nous

---

<sup>4</sup> Le corpus *Fidaplus*, corpus référentiel de la langue slovène en ligne, est composé de 600 millions de mots provenant des textes variés : la presse, la littérature, les textes juridiques et techniques. Il est accessible en ligne ([www.fidaplus.net](http://www.fidaplus.net)).

avons pu établir le taux de la grammaticalisation existante parmi les gérondifs et les participes, ce qui se confirme à travers la recherche dans le corpus *Fidaplus*, tandis que pour le français, nous avons préétabli la liste des gérondifs et participes grammaticalisés que nous avons testés dans les corpus parallèles bi- et plurilingues *FraSloK*<sup>5</sup> et *Evrokopus*<sup>6</sup>.

### 2.3.1 La recherche dans les corpus slovènes

Pour le slovène, une analyse des gérondifs en *-e*, gérondifs et participes en *-oč/-eč* et gérondifs en *-aje* a été effectuée à partir du corpus monolingue slovène *Fidaplus* ([www.fidaplus.net](http://www.fidaplus.net)).

La méthodologie du travail était la suivante : nous avons cherché dans le corpus tous les mots qui terminent en *-oč*, *-eč*, en nous servant de la consigne *-oč\**, *-eč\**<sup>7</sup>. Nous avons ensuite procédé à la désambiguïsation manuelle car les noms comme *pomoč*, *noč*, *tisoč*, *nekoč*, *moč* y figuraient à leur tour. La recherche était limitée à 100 000 occurrences. Les données ont été testées sur 3000 occurrences (1500 parmi les occurrences du début et 1500 à la fin de la liste des concordances données).

Les gérondifs et les participes en *-oč* les plus fréquents (plus de 2 dans 3000) étaient *rekoč* (*en disant*) (25 occurrences), *zahvaljujoč* (*grâce à*) (10), *takorekoč* (*soi-disant*, *pratiquement*) (9), *razumevajoč* (*comprenant*) (4), *upoštevajoč* (*considérant*) (3). Pour les gérondifs et les participes en *-eč*, nous avons obtenu *sodeč* (*jugeant*) et *obstoječ* (*existant*) (avec occurrences respectivement) qui avaient plus de 2 occurrences parmi les 3000 exemples. Ceci signifie que les gérondifs et les participes ne sont pas très fréquents en général, mais que quelques-uns parmi eux sont tout de même assez fréquemment employés.

L'analyse détaillée des gérondifs et participes choisis à travers tout le corpus *Fidaplus* a donné les résultats suivants. Le gérondif le plus fréquent était « *rekoč* » (33 241 occurrences) – en fonction d'adverbe de reformulation (exemple (11)), alors grammaticalisé :

- (11) *S Tomom sta tako rekoč ista generacija in danes oba letita pri Adrii Airways*  
*Avec Tomo, ils appartiennent pratiquement à la même generation et tous les deux sont*  
*pilotes chez Adria Airways.*  
(*Štajerski tednik 1996*)

<sup>5</sup> Le corpus parallèle français-slovène, compilé par Adriana Mezeg, contient 6 millions de mots et comporte des originaux et leurs traductions, y compris des textes littéraires et journalistiques.

<sup>6</sup> Le corpus composé par les textes législatifs de l'Union européenne, composé par le Service du gouvernement slovène de traduction et interprétation (SVEZ).

<sup>7</sup> L'astérisque signifie qu'il peut y avoir une suite – le concordancier fait le repérage de tous les mots qui ne sont pas au nominatif masculin, mais qui varient selon le genre, le nombre et le cas. Il s'agit surtout des participes présents adjectivalisés.

Dans 3221 cas, *tako rekoč* est écrit ensemble, ce qui présente un changement de la forme, une preuve de plus de sa grammaticalisation (exemple (12)) :

- (12) *Poznala sva **takorekoč** vse trgovske zviijače, upala sva si reči ne.*  
*Connaissant soi-disant toutes les ruses des commerçants, nous avons osé dire non.*  
*(Dolenjski list 1998)*

Un usage de *rekoč* comme noyau verbal a tout de même été trouvé (exemple (13)) :

- (13) *Pogosto se dogaja, da ljudje odkimajo z glavo, **rekoč**, veste, nisem jugonostalgik, ampak ...*  
*Il arrive souvent que les gens hochent la tête en disant vous savez, je ne suis pas nostalgique de la Yougoslavie, mais tout de même..*  
*(Delo 2004)*

Les gérondifs *zahvaljujoč* et *upoštevajoč* apparaissent dans leur rôle prépositionnel (exemples (14) et (15)) :

- (14) »*Zmago smo dosegli **zahvaljujoč** odlični obrambi. Samo 14 golov v naši mreži v 60 minutah je dokaz odgovornosti, ki so se je moji igralci še kako zavedali,«* je povedal trener Delmarja Bogosav Perić.  
*Nous avons gagné **grâce à** l' excellente défense – 14 buts en 60 minutes présentent la preuve de la responsabilité de nos joueurs, a dit/déclaré le coach de Delmar Bogosav Perić.*  
*(Dnevnik 1997).*
- (15) ***Upoštevaajoč** inflacijo to pomeni, da je dobiček realno manjši.*  
*L'inflation **prise en compte** cela signifie que le profit sera plus petit.*  
*(Delo 2003)*

Les formes en *-eč* retrouvées dans le corpus, à part *sodeč* et *obstoječ*, ne présentent que les participes adjectivalisés (les plus fréquents *rdeč* (*rouge*), *pereč* (*poignant*), (*ne*)*boleč* (*qui (ne fait pas) mal*), *svareč* (*qui annonce le mal*)) et les adverbes en *-eče* (*sledeče* (*suivant*), *klečeče* (*(s')agenouillant*), *dišeče* (*sentant*), *grozeče* (*menaçant*)), également les participes adjectivalisés qui ont subi l'adverbialisation. Parmi les 3000 occurrences des mots en *-eč\**, 14 exemples des formes en *-eč* ont été retrouvées avec un seul exemple du gérondif – *poti tihotapcev vodeč z vzhoda* (*les chemins des contrebandiers partant de l'est*) et il y avait plus d'une occurrence des gérondifs ou participes *sodeč* (6), *obstoječ* (5) et *boleč* (2 fois) et 2 restes de l'ancien participe lexicalisé *pereč*, *moreč*. Le gérondif le plus fréquent, *sodeč po* (*selon*), est grammaticalisé en tant que préposition. Dans le corpus Fidaplus, cette locution prépositionnelle atteint 13921 exemplaires (exemple (17)) :

(17) **Sodeč** po večrajšnji razpravi v kolegiju predsednika državnega zbora bo na njenem dnevnem redu okoli 20 točk.

**Vu / considérant / selon** le débat au conseil présidé par le président de l'Assemblée nationale hier, l'ordre du jour comportera 20 points environ.

(Dnevnik, 13. 2. 2002)

Nous pouvons conclure que les gérondifs les plus fréquents qui apparaissent en slovène avec la terminaison *-oč/-eč* sont ceux qui sont grammaticalisés.

La même procédure a été adoptée pour les gérondifs en *-aje* : dans les premières 3000 occurrences, nous avons trouvé 18 gérondifs différents, parmi lesquels 8 sont apparus plus d'une fois. Le résultat était presque identique pour les 3000 dernières occurrences parmi 100 000 occurrences avec la terminaison *-aje*. Les gérondifs plus fréquents *smehljaje* (*en riant*), *igraje* (*en jouant*), *šepetaje* (*en chuchotant*), tous dans l'usage adverbial, étaient confinés à l'apparence dans les oeuvres littéraires.

Dans l'usage prépositionnel, le gérondif le plus fréquent était *upoštevaje* avec 10 occurrences dans le test (et 1901 occurrences dans le corpus entier). Ce gérondif lexicalisé apparaît dans le corpus comme préposition *upoštevaje*, introduisant un nom à la forme accusative, ou une conjonction, plus exactement la locution conjonctive *upoštevaje, da -* (94 occurrences en corpus entier) que nous citons dans l'exemple (18) ci-dessous.

(18) **Upoštevaje** celoto je mogoče dopis koroškega deželnega glavarja z dne 21. decembra 2001 razumeti kot na nič stvarnega oprti poskus zunanjšega vplivanja na odločitve ustavnega sodišča.

**Prenant en compte** la totalité, la lettre du charintien du 21-décembre 2001 peut être comprise comme les tentations sans fondements de l'influence extérieure à la décision de la cour constitutionnelle ...

(Delo 2002)

Les formes gérondives en *-e*, étant déjà adverbialisées à partir du XVI<sup>e</sup> siècle, il était impossible de passer au repérage dans le corpus Fidaplus de la même manière que pour les formes précédentes. Nous avons analysé quelques exemples choisis : *molče* (*en silence*), *mimogrede* (*en passant*), *glede* (*selon*), parmi lesquels le dernier a subi une deuxième grammaticalisation en devenant la forme prépositionnelle *glede na* (*selon*).

### 2.3.2 L'analyse quantitative des gérondifs et participes français choisis dans deux corpus parallèles

Les corpus *FraSloK* et *Evroterm* sont composés d'une manière différente que Fidaplus et ne permettaient pas de recherche similaire à celle effectuée dans le corpus Fidaplus ; il nous a été impossible de recenser tous les gérondifs présents dans les deux corpus. Nous nous sommes donc limitées à des gérondifs et participes choisis pour lesquels nous savions déjà qu'ils avaient entamé le processus de grammaticalisation. Nous avons pris en compte les gérondifs et les participes suivants : (*en*) *admettant* (*que*), (*en*) *allant*, (*en*) *attendant* (*que*), *concernant*, *considérant* (*que*), *étant donné* (*que*).

Les résultats dans les deux corpus différaient considérablement. Le gérondif *en admettant* apparaissait dans le corpus *Evrokopus*<sup>8</sup> avec 27 occurrences, dans le corpus à valeur prépositionnelle (parmi lesquelles aussi l'exemple (18)). On compte 4 exemples où la forme participiale non grammaticalisée a été utilisée (entre autres dans l'exemple (19)) :

(18) *Ils ont confiance en ce système, tout en admettant la nécessité de poursuivre les travaux entrepris pour l'améliorer, notamment dans les régions où il existe un risque préoccupant de prolifération.*

(*Evrokopus* : ID: 21996A0520(01))

(19) (...) *pratiquant la vaccination contre la fièvre aphteuse et admettant sur son territoire la présence d'animaux vaccinés contre cette affection, ces animaux sont soumis aux exigences de la présente directive, exception faite de celles relatives à la vaccination antiaphteuse qui sont remplacées par les garanties suivantes...*

(*Evrokopus* : ID: 31977L0098)

Dans le corpus *FraSloK*<sup>9</sup>, *en admettant* n'avait que 8 occurrences, parmi lesquelles un exemple sous forme de conjonction ou locution conjonctive *en admettant que* (voir l'exemple (20)). Le reste des formes retrouvées était utilisé avec la valeur participiale ou gérondive.

(20) *On l'embrasse en arrivant et en partant, en **admettant** que « partir » ait un sens dans ce genre d'endroit (...).*

(*FraSloK*)

Le participe présent *allant* apparaît dans l'*Evrokopus* avec 601 occurrences et dans le corpus *FraSloK* avec 70 occurrences dont la plupart sont participiales ou gérondives (comme dans l'exemple (21)) et une plus petite partie prépositionnelle :

(21) *Je revis Charlotte dix ans après, pendant quelques heures, en **allant** à l'étranger. J'arrivai très tard le soir, et je devais repartir tôt ...*

(*FraSloK*)

L'utilisation prépositionnelle de l'exemple (22) est tirée d'*Evrokopus* :

---

<sup>8</sup> *Evrokopus* est le corpus parallèle des traductions de la législation européenne; sa partie française contient 25 millions de mots et 560 000 unités de traduction.

<sup>9</sup> Ce corpus parallèle, composé des originaux français du journal *Le Monde Diplomatique* ainsi que des textes littéraires et de leurs traductions slovènes, a été compilé par Adriana Mezeg dans le cadre de son travail de thèse doctorale et contient 6 millions de mots (Mezeg 2010).

- (22) *En cas de vacance simultanée des postes de Directeur général et de Vice-Directeur général, le Conseil d'administration élit, sur la base des candidatures reçues à la suite d'une mise au concours, un Vice-Directeur général pour la période **allant** jusqu'au prochain Congrès.*  
(Evrokopus : ID: Ur. l. RS, MP 2006-117)

Le gérondif et le participe (*en*) *attendant* apparaissent dans leurs formes grammaticalisées sous forme de la conjonction *en attendant que*, représentée par 361 occurrences dans l'*Evrokopus*, parmi lesquels l'exemple (23), dans la suite ou sous forme de la préposition *en attendant* (exemple (24)) :

- (23 (...) 6.23 *approuver, dans le cadre de ses compétences, les recommandations du Conseil d'exploitation postale concernant l'adoption, si nécessaire, d'une réglementation ou d'une nouvelle pratique en attendant que le Congrès décide en la matière ;*  
(Evrokopus : ID: Ur. l. RS, MP 2006-117)

- (24 (...) *au stockage des marchandises en dépôt temporaire (locaux fermés et emplacements, clôturés ou non, agréés par la douane, ou les marchandises sont stockées en attendant leur dédouanement) ;*  
(Evrokopus : ID: 21973A0518(01))

La recherche dans le corpus *FraSloK* a donné, en revanche, très peu d'exemples grammaticalisés parmi 136 gérondifs et participes recensés, l'exemple (25) ci-dessus présente *en attendant* dans son rôle adverbial :

- (25) *Espitalier va vous recevoir dans un quart d'heure, commença -t-il. En **attendant**, je vais vous expliquer pourquoi vous êtes là.*  
(*FraSloK*)

La forme grammaticalisée conjonctive *en attendant que* compte 12 occurrences dans *FraSloK* (exemple (26)) :

- (26) *(C'est) est la danse à laquelle tous les enfants du pays qui mangent des sardinelles **en attendant que** Dieu veuille bien penser à eux souhaiteraient pouvoir inviter leur mère...*  
(*FraSloK*)

Le participe présent *concernant* compte 17361 occurrences dans *Evrokopus*. A l'instar de l'exemple (27), tous avaient une valeur prépositionnelle.

- (27) (...) *du 22 février 2005 concernant les demandes d'adhésion à l'Union Européenne présentées par la République de Bulgarie et la Roumanie LA COMMISSION DES COMMUNAUTÉS EUROPÉENNES, vu le traité sur l'Union européenne, et notamment son article 49, considérant ce qui suit :*  
(Evrokopus : ID: UL EU, št. L 157)



Dans le corpus *FraSloK*, nous avons recensé 74 occurrences du participe présent *concernant*, la plupart avec la même valeur prépositionnelle (par exemple dans (28)), mais quelques-uns attestaient aussi un usage prédicatif (exemple (29)) :

(28) ... jamais elle ne laissait échapper un mot **concernant** son passé. Les notions évoquées plus haut d' héritage et de pension...

(*FraSloK*)

(29) (...) reçu , histoire de sauver la face. On échafaudait des plans insensés les **concernant**. Certains racontaient que j'allais emmener mon frère avec moi...

(*FraSloK*)

Pour le participe présent *considérant*, nous avons trouvé une nette distinction entre les données dans le corpus *FraSloK* et *Evrokopus*. Dans *FraSloK*, parmi les 22 occurrences de ce participe dans les textes littéraires, toutes les valeurs étaient prédictives (exemple (30)).

(30) (La) fille est devenue amorphe, elle est comme vidée de son énergie. Puis, la **considérant** avec des yeux embués de larmes, la vieille dame ajouta : – Je me doutais...

(*FraSloK*)

Les exemples tirés du journal *Le Monde Diplomatique* présentent les usages grammaticalisés (exemple (31)) :

(31) (...) un des rares dirigeants talibans au Sud-Waziristan, dénonça les accords, **considérant** que le Pakistan les avait violés . M. Tahir Yaldeshiv lui apporta aussitôt

(*FraSloK*)

La fréquence des formes grammaticalisées dans *Evrokopus* (23350 occurrences pour *considérant* tandis que 16318 pour *considérant que* ) est la preuve que ce mot fait partie du préambule de tout texte juridique – politique de l'UE (à comparer avec l'exemple (32) et (33)) :

(32) **Considérant** les conclusions du séminaire de Brighton de 1997 visant à encourager les projets pilotes d'autorisations uniques communautaires impliquant une coopération entre les administrations douanières des Etats membres ;

(*Evrokopus* : ID: Ur-l-RS-st-21-2009)

(33) (...) **considérant que** des études génétiques ont montré que les ours bruns des Pyrénées et les ours bruns de Slovénie appartiennent à la même lignée...

(*Evrokopus* : ID: Ur. l. RS, št. MP 2006-32)

Le participe présent composé *étant donné* est aussi, dans son rôle grammaticalisé de la conjonction *étant donné que*, l'indicateur des textes administratifs avec ses 914 exemplaires, parmi lesquels l'exemple (34) ci-dessous :

- (34) Par dérogation à l'article 5.2, El Salvador, le Panama (Rép.), les Philippines, la Rép. dém. du Congo et le Vénézuéla sont autorisés à ne pas renvoyer les colis après que le destinataire en a demandé le dédouanement, **étant donné** que leur législation douanière s'y oppose. (Evrokopus : ID: Ur. l. RS, MP 2006-117)

Sur la base de cette recherche dans les corpus de ces quelques participes et gérondifs grammaticalisés, nous pouvons démontrer que les formes grammaticalisées varient selon le type de texte dans lequel elles sont employées. Les recherches sur *Evrokopus* donnent les résultats qui diffèrent du *FraSloK*. Le nombre des gérondifs et participes grammaticalisés est plus élevé dans les textes administratifs et journalistiques, tandis que dans les textes littéraires, ceux-là gardent plus longtemps leur rôle de noyau de la prédication.

## CONCLUSION

Le procédé de la grammaticalisation, moteur du développement d'un côté de nouvelles formes grammaticales, et de l'autre, de nouvelles entrées lexicales, est fréquent aujourd'hui dans la classe des gérondifs et participes. Les gérondifs et les participes présents français aussi bien que slovènes présentent les cas de la grammaticalisation.

En français, la grammaticalisation des gérondifs et des participes présente des cas isolés, réservés à certaines formes, tandis qu'en slovène, la grammaticalisation peut toucher toute une catégorie des participes, notamment l'emploi adverbial des gérondifs en *-e*. En plus, en analysant l'apparition des participes et des gérondifs dans les corpus slovènes en ligne Fidaplus, nous avons pu constater que les « vrais » gérondifs et participes qui pourraient former le noyau verbal d'une structure détachée sont beaucoup plus rares que leurs homologues grammaticalisés.

Il y a une différence considérable concernant l'usage des formes impersonnelles des gérondifs ou participes français grammaticalisés dans le corpus *FraSloK* et *Evrokopus*. Dans le deuxième, l'usage grammaticalisé prévaut – tandis que dans le premier, nous trouvons encore les participes et les gérondifs dans leur usage comme noyau du syntagme verbal, surtout dans les textes littéraires. Sur la base de l'analyse des deux corpus, nous pouvons tirer la conclusion pour les textes français que la grammaticalisation est moins fréquente dans les textes littéraires qu'ailleurs. Nous pouvons exposer le facteur pragmatique ou textuel de la grammaticalisation des gérondifs et des participes.

Ces changements ne sont que partiellement notés par les dictionnaires, le dictionnaire du slovène standard *Slovar slovenskega knjižnega jezika* (SSKJ) n'en tient pas compte. Les grammaires mentionnent le procédé mais la recherche dans le corpus en donne de meilleures preuves dans des contextes particuliers. En revanche, le processus peut être remarqué dans les exemples des corpus mono- et bilingues étudiés. Les tendances sont visibles : les gérondifs et les participes slovènes les plus employés sont ceux qui ont subi la grammaticalisation. En revanche, en français, les cas isolés ont subi le figement prépositionnel, conjonctif ou adverbial tandis que le fonctionnement morpho-syntaxique et discursif de toute la catégorie des gérondifs et participes n'en est pas affecté.

## BIBLIOGRAPHIE :

- BLANCHE-BENVENISTE, Claire (1997) *Approches de l'oral en français*. Paris : Ophrys.
- BUSSMAN, Hadumond (1990) *Lexikon der Sprachwissenschaft*. Stuttgart : Kröner Verlag.
- FRUYT, Michèle (2009) « Adverbes latins, grammaticalisation et lexicalisation. » In : M. FRUYT/S. VAN LAER (éds.) *Adverbes et évolution linguistique en latin*. Paris : Kubaba (Série grammaire et linguistique). <http://www.editions-harmattan.fr/index.asp?navig=catalogue&obj=article&no=14024> accès avril-mai 2011
- HALMOY, Odile (2003) *Le gérondif en français*. Paris : Ophrys.
- JESENŠEK, Marko (1998) *Deležniki in deležja na -č in -ši*. Maribor : Zora.  
<http://www.ling.upenn.edu/~kroch/online.html> accès juillet 2011.
- LEHMANN, Charles (1985) « Grammaticalization : Synchronic Variation and Diachronic Change. » *Lingua e stile* 20, 303-318.
- LEHMANN, Charles (2005) « Theory and Method in Grammaticalization. » In: G. Diewald (éd.) *Grammatikalisierung*. Berlin : W. De Gruyter (Zeitschrift für Germanistische Linguistik, Themensheft).
- MARCHELLO-NIZIA, Christiane (2006) *Grammaticalisation et changement linguistique*. Bruxelles: De Boeck.
- MEILLET, Antoine. (1921, 1948<sup>2</sup>) *Linguistique historique et linguistique générale*. Paris, Honoré Champion.
- MEZEG, Adriana (2010) « Compiling and using a French-Slovenian parallel corpus. » In: R. Xiao (éd.) *Proceedings of The International Symposium on Using Corpora in Contrastive and Translation Studies : 2010 conference (UCCTS2010)*. Ormskirk : Edge Hill University.  
<http://www.lancs.ac.uk/fass/projects/corpus/UCCTS2010Proceedings/papers/Mezeg.pdf> accès avril-mai 2011.
- MOREL, Mary Annick/Laurent DANON-BOILEAU (1996) *Grammaire de l'intonation*. Paris : Ophrys.
- RAMOVŠ, Fran (1952) *Historična morfolologija slovenskega jezika*. Ljubljana : DZS.
- RIEGEL, Martin/Jean-Cristophe PELLAT/René RIOUL (1994) *Grammaire méthodique du français*. Paris : PUF.
- SCHLAMBERGER BREZAR, Mojca (2005) « Politična besedila kot tip besedil in postopki prevajanja stalnih formul.» In : Nike Kocijančič-Pokorn/Erich Prunč/Alessandra Riccardi (éds.) *Beyond equivalence*, (Graz translation studies, 9). Graz: Institut für Theoretische und Angewandte Translationswissenschaft, 121-135
- TOPORIŠIČ, Jože (1982) *Nova slovenska skladnja*. Ljubljana : DZS.
- TOPORIŠIČ, Jože (2000) *Slovenska slovnica*. Maribor : Obzorja.
- VINCENOT, Claude (1975) *Essai de grammaire slovène*. Ljubljana : Mladinska knjiga.
- WILMET, Marc (2007) *Grammaire rénovée du français*. Bruxelles : De Boeck.
- Dictionnaire de l'Académie en ligne <http://atilf.atilf.fr/dendien/>. Accès avril-juin 2011
- [http://www.lexilogos.com/francais\\_langue\\_dictionnaires.htm](http://www.lexilogos.com/francais_langue_dictionnaires.htm). Accès avril-juin 2011
- Evrokorpus, corpus parallèle français-slovène <http://evrokorpus.gov.si/>. Accès avril-juin 2011
- Fidaplus, corpus monolingue slovène [www.fidaplus.net](http://www.fidaplus.net). Accès avril-juin 2011
- FraSloK, corpus parallèle français-slovène <http://nl.ijs.si/spook/konkor/>. Accès avril-juin 2011
- Trésor de la langue française en ligne [www.atilf.atilf.fr](http://www.atilf.atilf.fr). Accès avril-juin 2011

## Résumé

### LE GÉRONDIF ET LE PARTICIPE PRÉSENT ET LEUR ÉVOLUTION VERS LA GRAMMATICALISATION : ÉTUDE CONTRASTIVE DU SLOVÈNE ET DU FRANÇAIS

Dans le cadre des études de la morphologie traditionnelle, le gérondif et le participe sont considérés être les formes verbales, formées à partir des désinences particulières, fait prouvé en diachronie. Ce procédé reste-t-il productif en synchronie ? A partir des fonctions exercées par les participes et les gérondifs dans le discours, nous pouvons nous poser plusieurs questions sur leur statut actuel aussi bien en français qu'en slovène. Ne s'agit-il pas du passage à la grammaticalisation en tant que procédé dans lequel les anciens gérondifs et participes endossent peu à peu le rôle d'une catégorie grammaticale dépendante ?

Sur ces bases, nous présenterons une étude contrastive sur la grammaticalisation des gérondifs et des participes dans les deux langues, menée sur deux corpus parallèles français-slovène accessible en ligne - *FraSloK* (<http://nl.ijs.si/spook/konkor/>) et *Evrokorpus* (<http://evrokorpus.gov.si/>), ainsi que sur le corpus monolingue slovène *Fidaplus* ([www.fidaplus.net](http://www.fidaplus.net)) qui mettra en valeur, à part les particularités morphologiques, aussi leurs particularités fonctionnelles dans leur usage contemporain. La recherche sur les gérondifs et les participes présents montrera que ces formes grammaticales ont suivi des procédés de grammaticalisation bien différents dans les deux langues. Tandis que les gérondifs et les participes restent productifs en français, surtout dans le registre littéraire, la grammaticalisation de ces formes est pratiquement généralisée en slovène, à quelques exceptions près, dans les textes administratifs et techniques.

## Povzetek

### DELEŽJA IN DELEŽNIKI SEDANJEGA ČASA IN NJIHOV RAZVOJ V SMERI GRAMATIKALIZACIJE: SLOVENSKO-FRANCOSKA KONTRASTIVNA ŠTUDIJA

V tradicionalni morfologiji so deležja in deležniki obravnavani kot glagolske oblike, ki se jih tvori s posebnimi končnicami. V diahroniji to dejstvo drži, vendar nas zanima, ali je ta pojav še produktiven v sinhroniji. Ob analizi funkcij, ki jih deležniki in deležja opravljajo v diskurzu, si lahko zastavimo več vprašanj o njihovem statusu v francoščini in slovenščini. Prihaja do pojava gramatikalizacije, pri katerem bivša deležja in deležniki v slovenščini in francoščini počasi prevzamejo vlogo odvisne slovnične kategorije. Na tej osnovi bomo predstavili kontrastivno študijo o gramatikalizaciji deležij in deležnikov v obeh jezikih, ki bo temeljila na analizi podatkov iz dveh vzporednih francosko-slovenskih korpusov, ki so dostopni prek spleta - *FraSloK* (<http://nl.ijs.si/spook/konkor/>) in *Evrokorpus* (<http://evrokorpus.gov.si/>) - ter enojezičnega slovenskega korpusa *Fidaplus* ([www.fidaplus.net](http://www.fidaplus.net)). Raziskava bo poleg oblikoslovnih značilnosti prikazala tudi funkcijske značilnosti deležnikov in deležij v njihovi sodobni rabi. S primeri bomo nakazali, da so te glagolske oblike sledile različnim procesom gramatikalizacije v vsakem od jezikov. V francoščini deležja in deležniki še ostajajo tvorni predvsem v literarnem jeziku, medtem ko je gramatikalizacija teh oblik v slovenskem jeziku takorekoč splošna - z izjemo besedil administrativnega in tehničnega značaja.

## JESPERSEN'S CYCLE IN SLOVENIAN

### 0. Introduction

The paper examines the syntactic properties of the negative marker in standard Slovenian and its Pannonian dialects in terms of Jespersen's cycle (Jespersen's 1924: 335). The term Jespersen's Cycle refers to the historic morpho-phonological process of grammaticalisation which involves repeated weakening and strengthening of the negative marker. Although the original claim (Jespersen *ibid.*) involves only the diachronic perspective, later research (Zanuttini 1991, Kiparsky and Corderavdi 2006, van der Auwera 2009 and *in press*, Willis *in press inter alii*) has shown that the same process can also be observed synchronically. In view of the latter claim, the paper argues that in modern Slovenian at least three morphologically distinct stages of Jespersen's cycle co-exist, namely, (i) the clitic-like negation, (ii) the adverb-like negation, and (iii) the bipartite negation. The syntactic account proposed herein assumes that there is a direct correlation between the morpho-phonological status of the negative marker and the derivation of negative clauses (cf.: Haegeman 1995). While a clitic-like negative marker heads the Negative Phrase and requires a syntactic host, an adverb-like negative marker occupies the specifier position of the Negative Phrase and functions as an independent syntactic constituent. In the case of Slovenian, the clitic-like negation is typical of standard Slovenian, while the adverb-like negation can be found in Pannonian dialects of Slovenian.

In addition, the paper explores the claim put forward by Zanuttini (1989) and Rowlett (1997) that the morpho-phonological status of the negative marker and its syntactic position(s) directly determine the semantic interpretation of multiple occurrences of negative elements. The authors suggest that the negative concord interpretation is possible only with a clitic-like negative marker, while the double negation interpretation is typical of languages with an adverb-like negative marker. The analysis of the relevant Slovenian data disproves the claim and shows that there is no direct correlation between the morpho-phonological status of the negative marker and the semantic interpretation of multiple occurrences of negative elements.

The paper is organized as follows. Section 1 offers a concise outline of the theoretical framework which forms the background to this paper. Section 2 deals with the negation and the status of the negative marker in standard Slovenian (2.1) and the Pannonian dialect group (2.2). In section 3, we present the relevant data, and

---

\* *Author's address:* Filozofska fakulteta, Oddelek za anglistiko in amerikanistiko, Aškerčeva 2, 1000 Ljubljana. Email: gasper.ilc@ff.uni-lj.si

develop a generative syntactic account of the Slovenian negative marker in terms of Jespersen's cycle. Section 4 concludes the paper.

## 1 THEORETICAL BACKGROUND

### 1.1 Jespersen's cycle

It is a well-known fact that in many languages the negative marker has historically undergone the cyclic process of morpho-phonological weakening and strengthening known as Jespersen's cycle. The term was first used by Dahl in his seminal paper (1979) on the typology of negation, and it is accredited to Jespersen's (1924: 335) observation about the diachronic status of negative markers in sundry languages:

The negative adverb is often weakly stressed, because some other word in the sentence has to receive a strong stress of contrast. But when the negative has become a mere proclitic syllable or even a single sound, it is felt to be too weak, and has to be strengthened by some additional word, and this in its turn may come to be felt as the negative proper, which then may be subject to the same development as the original word. We have thus a constant interplay of weakening and strengthening, which with the further tendency to place the negative in the beginning of the sentence where it is likely to be dropped [...] leads to curious results.

Jespersen (1924: 335)

Jespersen (*ibid.*) identifies five different stages of the described cycle:

**Stage 1.** Negation is expressed by a pre-verbal negative marker, typically associated with the morpho-syntactic properties of a clitic (1a).

**Stage 2.** Negation is expressed by a bipartite negative marker, consisting of a weak pre-verbal element and a reinforcing post-verbal negative element. The reinforcing element acquires a secondary grammatical meaning even while retaining some of its original lexical meaning (1b). At this stage, the grammaticalisation process of the reinforcing element has begun.

**Stage 3.** The original pre-verbal marker is reduced to a phonologically null element; the post-verbal negative marker is completely grammaticalised, and functions as the sole negative marker (1c).

**Stage 4.** The post-verbal negative marker moves to the pre-verbal position (1d), but still retains some of its adverb-like syntactic properties.

**Stage 5.** The new pre-verbal negative marker becomes subject to a new process of weakening (1e). The cycle has come full circle: stages 1 and 5 display the same morpho-syntactic properties of the negative marker.

In what follows, we will use the term clitic-like negation for stages 1 and 5, bipartite negation for stage 2 and adverb-like negation for stages 3 and 4.<sup>1</sup>

- (1) a) *ic ne secge.*                      English (diachronic perspective, from Jespersen 1924: 335-6)  
b) *I ne seye not.*  
c) *I say not.*  
d) *I do not say.*  
e) *I don't say.*  
'I don't say.'

Although Jespersen (*ibid.*) builds his argumentation from the diachronic perspective, more recent syntactic analyses have argued (Haegeman 1995, Zanuttini 1997 *inter alii*) that within the same language different stages of Jespersen's cycle can be found synchronically. In French, for instance, stage 1 is still present in literary or elevated contexts with the expletive or pleonastic negation (2a) and some negated modals (2b). Stage 2, the bipartite negation, is commonly used in standard French (2c), while stage 3 is typical of colloquial and Québécois French.

- (2) a) *J'ai peur qu'elle ne soit déjà là.*  
I have fear that-she not is-subj. already there  
'I'm afraid that she might already be here.'
- b) *Je ne savais comment répondre.*  
I not knew how to-reply  
'I didn't know how to reply.'
- c) *Je ne parle pas chinois.*  
I not speak not Chinese  
'I don't speak Chinese.'
- d) *Je sais pas quoi faire.*  
I know not what to-do  
'I don't know what to do.'

It is noteworthy that in recent linguistic theory, the term Jespersen's cycle is no longer associated only with negation, but refers to any process of grammaticalisation, during which an independent lexical item gradually acquires a secondary grammatical meaning and function, and is consequently subject to a morpho-phonologi-

---

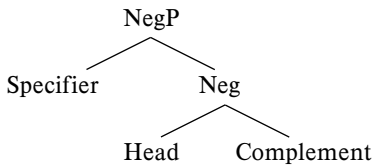
<sup>1</sup> For a detailed presentation of the stages within Jespersen's cycle and cross-linguistic typology see *inter alii* Kiparsky and Corderavdi 2006, van der Auwera 2009 and *in press*, and Willis *in press*.

cal weakening, the final result of which is a reduction to a clitic, an affix or even a phonologically null element (cf.: Van Gelderen 2004 *inter alii*).<sup>2</sup>

**1.2 Generative grammar, negation and Jespersen’s cycle: basic tenets**

Ever since Pollock’s (1989) influential work on the verb movement, generative analysis has assumed (Laka 1990, Haegeman 1995, Cinque 1999 *inter alii*) that negated clauses contain an additional functional projection, the Negation Phrase (*henceforth*: NegP), which according to the X-bar theory of phrase structure<sup>3</sup> comprises of the head, the specifier and the complement (3a).

(3) a) the structure of the Negation Phrase (NegP):



With regard to the hierarchical order of the functional projections, Chomsky (1998: 15) proposes the following system of core functional categories: (i) C(omplementizer) expressing force/mood, (ii) T(ense) expressing tense/event structure, and (iii) *v* (light verb) heading the verbal (transitive) constructions. The hierarchical ordering of these functional categories depends on their selectional properties: C selects T, while T selects *v*, which in turn selects verbal (V) elements (3b). It needs to be stressed, however, that Chomsky (op. cit.: 15, fn 31) uses functional categories C and T “as surrogates for richer systems”, such as those proposed by Pollock (1989), Rizzi (1997) and Cinque (1999); therefore, the analysis in the subsequent sections assumes the presence of at least one additional functional category, Neg(ation), in negative clauses. The functional category Neg is selected by T (3c).<sup>4</sup>

<sup>2</sup> For example, in Slavic languages some locative prepositions have undergone the process of grammaticalisation, during which they have been reduced to verbal prefixes denoting the perfective aspect:

- (i) *Janez je skočil prek ovire.* Slovenian  
Janez is jumped over hurdle
- (ii) *Janez je preskočil oviro.* Slovenian  
Janez is over-jumped hurdle  
'Janez jumped over the hurdle.'

<sup>3</sup> Since the 1970s X-bar theory has been standardly assumed in Chomskian grammar, including early work on Minimalism (cf. Chomsky 1995, chapters 1-3). More recently, however, Chomsky has proposed to eliminate the X-bar theory as a separate module of the grammar, arguing that restrictions on the form of structural descriptions follow directly from the properties of structure-building processes themselves (cf. op. cit.: chapter 4, Chomsky 1998, 1999, 2001).

<sup>4</sup> Both categories, T and Neg, are typically associated with the tense/event structure (cf.: Cinque 1999).



(3) b) clausal structure (non-negative sentences):

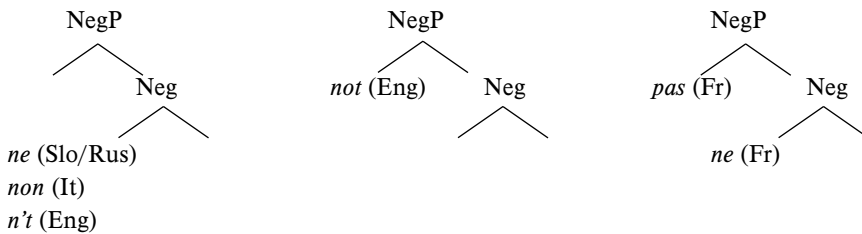
CP [TP [vP [vP ]]]

c) clausal structure (negative sentences):

CP [TP [NegP [vP [vP ]]]]

The syntactic accounts of Romance (Zanuttini 1989, Laka 1990 *inter alii*), Germanic (Haegeman 1995 *inter alii*), and Slavic (Abels 2005, Sue Brown and Adam Przepiorkowski 2007, Ilc and Milojević Sheppard 2005 *inter alii*) languages, have claimed that there is a direct correlation between Jespersen's Cycle and the structural position of the negative marker within the NegP. In languages with a weak, clitic-like preverbal negative marker (stages 1 and 5), the negative marker is in the head position of the NegP. In languages with a strong, adverb-like negative marker (stages 3 and 4), the negative marker occupies the specifier position of the NegP. In the case of a bipartite negative marker (stage 2), the weakened element is in the head position, and the reinforcing element in the specifier position:

(4) a) weak negative marker:      b) strong negative marker:      c) bipartite negative marker:



Zanuttini (1989) and Rowlett (1997) explore the relation further, and claim that there is a direct correlation between the syntactic position of the negative marker (4) and the semantic interpretation of multiple occurrences of negative elements, more precisely, the negative concord interpretation and the double negation interpretation. The term negative concord (*henceforth*: NC) refers to all cases in which multiple occurrences of morphologically negative elements are interpreted as a single semantic negation (4d). In contrast, the double negation interpretation (*henceforth*: DN) follows the logical reasoning in which one negation cancels the other (4e).

(4) d) *Je n'ai vu personne.* (NC) French  
 I not-have seen nobody  
 'I didn't see anybody.'

e) *I did not see nobody.* (DN) standard English  
 'I saw somebody.'

Zanuttini (1989) and Rowlett (1997) suggest that NC is possible only in languages in which the negative marker occupies the head position, for example, in Romance languages. On the other hand, the DN exists in languages with the negative marker in the specifier position, typically in the Germanic language group. This claim, however, is refuted by Haegeman (1995):

[T]here is no general correlation between NC and the presence of an overt Neg [head]. The Bavarian dialect of German [...] lacks an overt negative head and yet it has NC (Bayer 1990); in WF and in French, NC available regardless of the overt realization of the negative head.

Haegeman (1995: 165-166)

In section 2.3, the analysis of the Slovenian data confirms Hageman’s (1995) conclusion.

## 2 NEGATION IN SLOVENIAN

### 2.1 Standard Slovenian

In standard Slovenian, negation is expressed with the particle *ne*, which is typically unstressed in neutral and/or unmarked environments.<sup>5</sup>

The negative marker *ne* displays characteristics typical of a verbal proclitic, requiring a strict adjacent position to the finite-verb (5a). Toporišič (2000: 671) claims that the particle *ne* is the final element of the Wackernagel clitic cluster (*i.e.* clitic-second position); this, however, is not a valid assumption since the particle *ne* can occur outside the Wackernagel clitic cluster (Orešnik 1985-86: 213), and can together with the finite verb function as a host to the Wackernagel clitic cluster (5b), as argued by Milojević Sheppard and Golden (2000: 96). These facts lead to the conclusion that the combination *ne+finite verb* functions as a single syntactic constituent, as in (5b). In the case of the present tense forms of the verbs *biti* (to be), *imeti* (to have) and *hoteti* (to want), the two elements form a morpho-phonological constituent: *nisem*,

<sup>5</sup> The occurrences of the stressed negative marker are rare and/or contextually specific, for example, with the non-finite verbs (i), VP preposing (ii), and VP ellipsis (iii).

(i) *Ne se mi več opravičevati!* Golden 2000: 43, (74d)  
 not refl.cl. to-me-cl. anymore apologise  
 ‘Stop apologising to me.’

(ii) *Dam toliko ne, posodim pa.* Golden 2000: 44, (79)  
 give so-much not lend but  
 ‘I won’t give so much, but I can lend it.’

(iii) *Janez ne bo prišel in tudi Marija verjetno ne.*  
 Janez not will come and also Marija probably not  
 ‘John won’t come and also Mary won’t.’

Since the negative marker *ne* in the stressed position is no longer a clitic, it does not require a verbal host, therefore, it functions syntactically as an independent constituent.

*nimam* and *nočem* respectively (5c). In question formation, the tense auxiliary must move to the complementizer position together with the negative marker (5d).

(5) a) *Janez ne (\*nikoli/rad) bere rumenega tiska.*  
 Janez not never/glad read yellow press  
 ‘Janez doesn’t read yellow press.’

b) [*Ne dam*] *ti ga.*  
 not give you-cl. it-cl.  
 ‘I won’t give it to you.’

c) *Janez ni bral rumenega tiska.*  
 Janez not+is read yellow press  
 ‘Janez didn’t read yellow press.’

d) [<sub>CP</sub> [<sub>C</sub> *Ali nisi*]<sub>i</sub>] [<sub>T</sub> *t*]<sub>i</sub>] [<sub>NPeg</sub> *t*]<sub>i</sub>] [<sub>VP</sub> *videla Micka?*]]<sup>6</sup>  
 Q not-are seen Micka  
 ‘Haven’t you seen Micka?’

In regard to Jespersen’s cycle, Slovenian examples with the unstressed negative marker *ne* (5a-d) place standard Slovenian at stage 1 of the cycle: negation is expressed by a pre-verbal negative marker with the morpho-syntactic status of a clitic.

Slovenian, like other Slavic languages, is a typical negative concord language: the overt negative marker *ne* and the co-occurring negative elements are interpreted as a single negation (6a). In terms of its scope (Ilc *in press*),<sup>7</sup> the negative marker *ne* is limited to a finite clause and its non-finite complements (6b), unless introduced by a *wh*-word (6c). The negative marker in the matrix clause cannot have scope over a finite complement clause (6d).

(6) a) *Nihče ne bo odšel nikamor.*  
 nobody not will gone nowhere  
 ‘Nobody will go nowhere.’

<sup>6</sup> Cf.: (i) \**Ali si ne videla Micka?*  
 Q are not seen Micka  
 ‘Haven’t you seen Micka?’

<sup>7</sup> The scope of the negative marker in (standard) Slovenian can be easily determined by the presence of the Genitive of negation: a direct object within the scope of negation is marked with the Genitive case instead of the Accusative case.

- b) *Učitelj ne sili študentov reševati takšnih problemov.*  
 teacher not forces students-GEN solve-INF such problems-GEN  
 ‘The teacher does not force the students to solve such problems.’
- c) *\*Ne vem zakaj napisati pisma.*  
 not know why write letter-GEN  
 ‘I don’t know why to write a letter.’
- d) *\*Janez ne pravi, da bere časopisov.*  
 Janez not says that reads newspapers-GEN  
 ‘John doesn’t say that he reads newspapers.’

## 2.2 Pannonian dialect group

The Pannonian dialect group is spoken in the North-eastern parts of Slovenia, including Prekmurje and the Easternmost areas of Styria, and is comprised of four dialects: (i) Prekmurje dialect, (ii) Slovenian Hills dialect, (iii) Prlekija dialect and (iv) Haloze dialect (Smole 1998: 1-5).

### 2.2.1 Data

In the Pannonian dialect group, the negative marker appears in two morphologically related forms. First, the adverb-like negative marker *ne* (also pronounced in some dialects as *nei*) does not require the adjacent position to the finite verb (7a-e), is not typically merged with the present tense form of the auxiliary *to be* (7a-c), and can independently function as a host to a Wackernagel clitic (7e). If the finite verb is the tense auxiliary (7a,b) or the copula (7c), the negative marker follows the verbal form, but if the finite verb is a full lexical verb, then the negation precedes the verbal form (7d,e). In questions, the tense auxiliary moves to the complementizer position without the negative marker (7f), which is not the case with the standard Slovenian (5d). In the Prekmurje dialect, the present form of the auxiliary *to be* and the negative marker *ne* are merged (7g,h), but only if they are stressed (Zorko 1998: 232).

- (7) a) *Še plačati je mugo, pa ga je nikdar ne več vzeo na delo.*  
 even paid is had-to but him is never not anymore taken on work  
 ‘Even though he had to pay for it, he was never hired for work anymore.’
- b) *Počitka smo ne meli kak moja zeti tota deca.*  
 rest are not had as have now these children  
 ‘We didn’t have so much rest as children have today.’
- c) *To je ne tak.*  
 this is not so  
 ‘This isn’t so.’

d) *Tega nene dobro vem.*  
 this not well know  
 'I don't know this very well.'

e) *Ne me izdoti.*  
 not me betray  
 'Don't betray me.'

f) [CP[C *Si*<sub>i</sub> *ga* [T<sup>t</sup><sub>i</sub>] [NegP *nei* [Neg ] [vP *čula?*]]]  
 are him not heard  
 'Have you not heard him?'

g) *Pa jes san tou nei poznala, pa neisan znala.*  
 but I am this not known but not-am known  
 'Since I wasn't familiar with it, I didn't know how to do it.'

h) *Nikoga pa nei takšnoga bilou, ka bi se gor postavu.*  
 nobody not-is that been who would self up stand  
 'There was nobody who would stand up for it.'

Second, the negative particle also occurs in the form *nene*, which can be seen as a result of the morpho-phonological reduplication. Same as *ne/nei*, the form *nene* is typically stressed in unmarked and/or neutral environments, therefore, identical syntactic features can be observed:

(8) a) *Nene te poznan.*  
 not you know  
 'I don't know you.'

b) *Tega sn še nene čula.*  
 This am yet not heard  
 'I haven't heard that before.'

Occasionally, though not very frequently, a bipartite negative marker can be found: a weak clitic-like negative marker is merged with the finite verbal form, and then reinforced by the adverb-like marker *ne/nei*:<sup>8</sup>

<sup>8</sup> It is unclear at this point whether the negative marker *nei* in (9) functions as the reinforcing negative element of the bipartite negative marker *ne...nei* or whether it is a marker of negative polarity in answers to questions (cf.: *No, we do not know*). Explanations provided by Zorko (2009: 230) and judgements by some speakers of Pannonian dialects suggest that the former explanation is most promising.

(9) a) *Nei ne nūcamo.*  
not not need  
'We don't need it.'

b) *Nei ne veimo.*  
not not know  
'We don't know.'

As in standard Slovenian, multiple occurrences of negative constituents are typically interpreted as the negative concord in the Pannonian dialect group:

(10) *Nikaj je nei.*  
nothing is not.  
'There's nothing.'

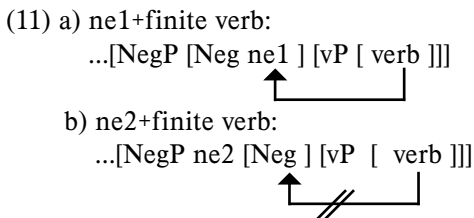
### 3 DATA DISCUSSION AND ANALYSIS

In this section, we develop a syntactic analysis of the negative marker both in the standard Slovenian and in the Pannonian dialect group in terms of Jespersen's cycle. Based on the data presented and discussed in section 2, we argue that in standard Slovenian the unstressed negative marker *ne* (henceforth:  $ne_1$ ) in (5a-d) should be analysed as the head of the NegP. Due to its relatively weak syntactic properties (stage 1 of Jespersen's cycle),  $ne_1$  requires that it be merged with the finite verbal form during the derivation. This syntactic step is achieved by the movement of the finite verb from its base-generated position in the verbal layer of the clause structure (3b) to the head position of the NegP (3a,c). As soon as the two elements are merged, the unit  $ne_1$ +finite verb is accessible to the derivation only as a single syntactic constituent. This fact is best captured by (5b), where the unit  $ne_1$ +finite verb functions as a host to a Wackernagel clitic, and by (5c), where the unit  $ne_1$ +finite verb moves overtly to the complementizer position. In the Pannonian dialect group, there is a homophonic negative marker *ne* (also spelt-out as *nei* or *nene*; henceforth:  $ne_2$ ) with adverb-like properties, which occupies the specifier position of the NegP. Since  $ne_2$  is not clitic in nature, it does not require a verbal host and is thus accessible to the derivation as an independent syntactic unit; for example, in (7e) it hosts a Wackernagel clitic, and in questions it remains *in situ* when the tense auxiliary moves to the complementizer position (7f). Examples (9) show that in the Pannonian dialect group, both  $ne_1$  and  $ne_2$  can co-occur in a form of a bipartite negative marker (stage 2 of Jespersen's cycle). Thus, we can assume that  $ne_1$  has undergone the process of a morpho-phonological weakening, and has been reinforced by a homophonic adverb-like lexical element  $ne_2$ . The different syntactic properties of the two negative markers can be observed in (9): the clitic-like  $ne_1$  is adjacent to the finite verb, whereas  $ne_2$  in the sentence initial position is separated from the finite verbal form. Since the occurrences of the bipartite negative marker  $ne_1$ ... $ne_2$  in the Pannonian dialect group are scarce, we can conclude that the morpho-syntactic weakening of  $ne_1$  has reached its

endpoint, the result of which is a phonologically null element, leaving the reinforcing negative marker  $ne_2$  as the sole negative marker (stages 3 and 4 of Jespersen's cycle). It is on account of its observed syntactic properties in (7) and (8) that we propose it be analysed as the specifier of the NegP.

The asymmetric position of the negative marker  $ne_2$  with regard to the finite verb (7c,d) can be explained in terms of the structural position of the finite verb. Assuming the hierarchical order of the functional projections to be T>Neg> $v$  (3), and the fact that full lexical verbs are base-generated in the verbal layer of functional projections, then  $ne_2$  in the specifier position of the NegP must precede the finite verb in  $v$  (Neg> $v$ ). The auxiliaries and the copular verbs in Slovenian, however, move overtly to T (Ilc and Sheppard 2003), so they must precede  $ne_2$ , which is in a structurally lower position (T>Neg).

Comparing now the derivation of sentences with  $ne_1$  and  $ne_2$  with regard to the position of the finite verb, the proposal claims that there are two different derivations. In the case of the  $ne_1$ , the finite verb must move from its base-generated position to Neg to be merged with the clitic negative marker (11a). As soon as they merge, they are accessible to further derivation only as a syntactic constituent (5a). The marker  $ne_2$ , on the other hand, lacks the clitic nature, so the finite verb does not move to the Neg to host the clitic. As a consequence, the finite verb remains *in-situ* (11b). The fact that  $ne_2$  and the finite verb do not form a syntactic unit can best be observed in (7a,d), where  $ne_2$  is separated from the finite verb by *nikdar* and *dobro* respectively (cf.: (5a)). The negative marker is also unaffected by the movement of the tense auxiliary to the complementizer position (cf.: (5d) and (7f)).



To sum up, diagram 1 represents the stages of Jespersen's cycle in Slovenian from a diachronic perspective.

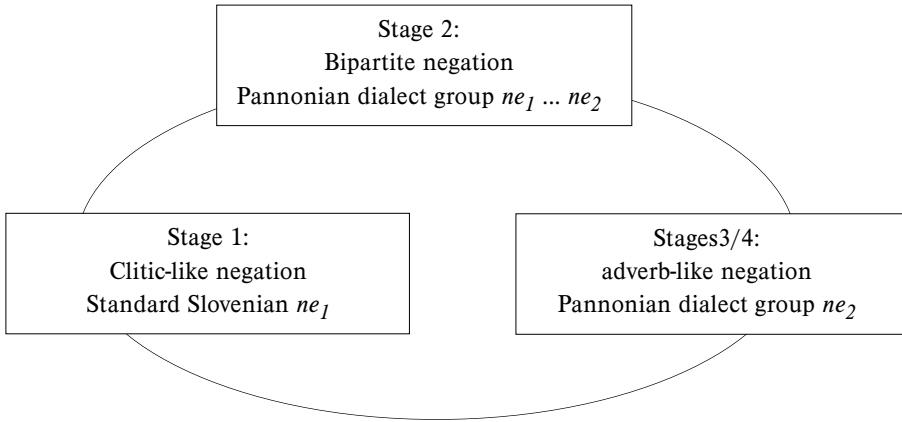


Diagram 1: Jespersen's cycle in Slovenian from a diachronic perspective.

What remains to be addressed is the syntactic difference between the emphatic use of the unit  $ne_2 + to\ be_{(present)\ tense} - neisan$  and  $nei$  in (7g,h) – and their non-adjacent counterparts  $san\dots nei$  and  $je\dots nei$ . Speakers' judgements with regard to their usage as well as their (non-) emphatic status vary, so it remains unclear whether the two forms should be treated as interchangeable or not. Hence, the question remains unanswered, and merits further research into the subject matter.

Examples containing the negative marker as well as negative elements ((7a,h) and (10)) demonstrate that regardless of the syntactic position of the negative marker, be it in the head (stages 1 and/or 2) or the specifier position (stages 2-4), multiple occurrences of negative elements in Slovenian always give rise to the negative concord interpretation. Slovenian examples thus confirm the conclusion by Haegeman (1995: 165-166, cf.: section 1.2) that there is no general correlation between NC/DN and the presence of an expressed element in the Neg.

#### 4 CONCLUSION

In this paper, we have argued that, synchronically, Slovenian displays at least three stages of Jespersen's cycle: (i) the clitic-like negation in Standard Slovenian (stage 1), (ii) the bipartite negation in the Pannonian dialect group (stage 2), and the adverb-like negation in the Pannonian dialect group (stages 3 and 4). For the clitic-like negation we have proposed the analysis which places the negative marker in the head position of the NegP. In addition, due to its proclitic nature, it is merged with the finite verbal form into a single syntactic constituent. In the case of the adverb-like negation, the negative marker occupies the specifier position of the NegP, and behaves as an independent syntactic constituent. The bipartite negation combines the properties of both the clitic-like and adverb-like negations: the weak negative



marker occupies the head position, whereas the stronger negative marker occupies the specifier position within the NegP.

With regard to the possible interpretations of multiple occurrences of the negative elements, there is no direct correlation between the morpho-syntactic properties of the negative marker and the negative concord or double negation interpretation, since in Slovenian, the negative concord interpretation is available in all cases.

### **Bibliography**

- ABELS, Klaus (2005) "Expletive negation' in Russian: A conspiracy theory." *Journal of Slavic Linguistics* 13, 5-74.
- BROWN, Sue/Adam PRZEPIORKOWSKI (eds.) (2007) *Negation in Slavic*. Bloomington. In: Slavica Publishers.
- CARDINALETTI, Anna/Maria Teresa GUASTI (1993) "Negation in epistemic small clauses." *Probus* 5, 39-61.
- CHOMSKY, Noam (1995) *The Minimalist Program*. Cambridge, MA: The MIT press.
- CHOMSKY, Noam (1998) "Minimalist Inquiries: the Framework." *MIT Occasional Papers in Linguistics* 15. Cambridge, MA: The MIT press.
- CHOMSKY, Noam (1999) "Derivation by phase." *MIT Occasional Papers in Linguistics* 18. Cambridge, MA: The MIT press.
- CHOMSKY, Noam (2001) "Beyond Explanatory Adequacy." *MIT Occasional Papers in Linguistics* 20. Cambridge, MA: The MIT press.
- CINQUE, Guglielmo (1999) *Adverbs and Functional Heads. A cross-Linguistic Perspective*. Oxford: Oxford University Press.
- DAHL, Östen (1979) "Typology of sentence negation." *Linguistics* 17, 79-106.
- GELDEREN, Elly van (2004) *Grammaticalization as Economy*. Amsterdam: John Benjamins.
- GOLDEN, Marija (2000) *Teorija opisnega jezikoslovja. 1. skladnja*. Ljubljana: Filozofska fakulteta.
- GOLDEN, Marija/Milena MILOJEVIĆ SHEPPARD (2000) "Slovene pronominal clitics." In: F. Beukema/M. den Dikken (eds.) *Clitic Phenomena in European Languages*. Amsterdam/Philadelphia: John Benjamins, 191-207.
- HAEGEMAN, Liliane (1995) *The syntax of negation*. Cambridge: CUP.
- ILC Gašper/Milena MILOJEVIĆ SHEPPARD (2005) "Negative pronouns in Slovene : a minimalist account." *Linguistica* 45, 107-120.
- ILC, Gašper (*in press*) "Optionality of the Genitive (of Negation) in Slovene." Proceedings of FDSL 8, Potsdam 2009.
- ILC, Gašper/Milena MILOJEVIĆ SHEPPARD (2003) "Verb movement in Slovenian : a comparative perspective." In: J. Orešnik/D. F. Reindl (eds.). *Slovenian from a typological perspective*, (Sprachtypologie und Universalienforschung = Language typology and Universals, 56(3)). Berlin: Akademie Verlag, 266-286.
- JESPERSEN, Otto (1924) *The Philosophy of Grammar*. London: George Allen&Unwin Ltd.
- KIPARSKY, Paul/Cleo CONDORAVDI (2006) "Tracking Jespersen's cycle." In: M. Janse/B.D. Joseph/A. Ralli (eds.) *Proceedings of the 2nd International Conference of Modern Greek Dialects and Linguistic Theory*. Mytilene: Doukas, 172-197.

- LAKA, Miren Itziar Mugarza (1990) "Negation in syntax: on the nature of functional categories and projections." PhD Thesis. MIT: Department of Linguistics and Philosophy.
- OREŠNIK, Janez (1985-86) "O naslonskem nizu v knjižnji slovenščini." *Jezik in slovstvo* 31, 213-215.
- ROWLETT, Paul (1997) "Jespersen, Negative Concord and A<sup>2</sup>-Binding." In: D. Forget/P. Hirschbühler/F. Martineau/M. L. Rivero (eds.) *Negation and Polarity, Syntax and Semantics*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 323-340.
- SMOLE, Vera (1998) "Slovenska narečja" *Enciklopedija Slovenije* 12, 1-5, Ljubljana: Mladinska knjiga.
- TOPORIŠIČ, Jože (2000) *Slovenska Slovnica*. Maribor: Obzorja.
- VAN DER AUWERA, Johan (2009) "The Jespersen cycles." In: E. van Gelderen (ed.) *Cyclical change*. Amsterdam: Benjamins, 35-71.
- VAN DER AUWERA, Johan/L. Van Alsenoy (in press) "Indefinite pronouns, synchrony and diachrony - comments on Willis." In: P. Larrivière/R. Ingham (eds.) *The evolution of negation: Beyond the Jespersen cycle*. Berlin: Mouton de Gruyter.
- WILLIS, David (in press.) "A minimalist approach to Jespersen's Cycle in Welsh." In: Jonas, D./J. Whitman/A. Garrett (eds.) *Grammatical Change: Origins, Nature, Outcomes*. Oxford: Oxford University Press.
- ZANUTTINI, Raffaella (1991) *Negation and Clausal Structure. A comparative Study of Romance Languages*. Oxford: OUP.
- ZANUTTINI, Raffaella (1989) "The structure of Negative clauses in Romance." Paper presented at the XII GLOW conference, Utrecht.
- ZORKO, Zinka (1998) *Haloško narečje in druge dialektološke študije*. Maribor: Dravska tiskarna.
- ZORKO, Zinka (2009) *Narečjeslovne razprave o koroških, štajerskih in panonskih govorih*. Maribor: Dravska tiskarna.

#### Abstract

#### JESPERSEN'S CYCLE IN SLOVENIAN

The paper examines the syntactic status of the negative marker in standard Slovenian and its Pannonian dialects in terms of the grammaticalisation process known as Jespersen's cycle. Assuming that Jespersen's Cycle can be observed synchronically, the paper focusses on the correlation between the morpho-phonological strength of the negative marker and the syntactic derivation of negative clauses. The data analysis identifies at least three different stages of Jespersen's cycle in modern Slovenian: (i) the clitic-like negation, (ii) the bipartite negation, and (iii) the adverb-like negation, the first occurring in standard Slovenian and the latter two in the Pannonian dialect group. In terms of the generative syntactic derivation, the analysis proposes that the negative marker occupies three different structural positions: (i) the head of the Negation Phrase (clitic-like negation), (ii) the specifier of the Negation phrase (adverb-like negation) or (iii) both syntactic positions (bipartite negation). In addition, the paper explores the question whether the syntactic position of the negative marker determines the semantic interpretation of multiple occurrences of negative elements, in particular, the neg-

ative concord and the double negation interpretation. The analysis shows that in Slovenian the morpho-phonological properties of the negative marker and its structural position bear no consequences for the semantic interpretation of multiple occurrences of negative elements.

Povzetek

### JESPERSENOV KROG V SLOVENŠČINI

Prispevek obravnava skladenjske lastnosti označevalca zanikanja v knjižni slovenščini in v panonski narečni skupini z vidika procesa morfofonološkega krčenja, v literaturi znanega kot Jespersenov krog. Izhajajoč iz uveljavljene teze, da lahko različne razvojne stopnje Jespersenovega kroga opazujemo tudi sinhrono, se prispevek osredini na soodvisnost med morfofonološkimi lastnostmi nikalnega označevalca in skladenjsko izgradnjo nikalnih stavkov. Analiza podatkov pokaže, da lahko v sodobni slovenščini govorimo o vsaj treh različnih stopnjah Jespersenovega kroga, in sicer o naslonskem, prislovnem in dvodelnem zanikanju. Prvo je značilno za knjižno slovenščino, medtem ko moremo slednja najti v panonski narečni skupini. Na osnovi pridobljenih podatkov prispevek razvije razčlemba skladenjske izgradnje nikalnih stavkov, ki temelji na sodobnem tvorbeno-pretvorbenem pristopu. Razčlemba pokaže, da lahko označevalec zanikanja zaseda dve različni skladenjski mesti v nikalni zvezi, in sicer jedro pri naslonskem tipu zanikanja, določilo pri prislovnem tipu zanikanja ter obe mesti hkrati pri dvodelnem tipu zanikanja. Prispevek pokaže, da morfofonološke lastnosti označevalca zanikanja vplivajo samo na skladenjsko izpeljavo nikalnih stavkov, medtem ko nimajo vpliva na njihovo semantično interpretacijo.



## QU'EST-CE QUE LE SLOVÈNE ET LE FRANÇAIS ONT EN COMMUN ? LES ARTICLES DÉFINI ET INDÉFINI

### 0. INTRODUCTION

La catégorie du défini et de l'indéfini occupe un chapitre à part dans la grammaire slovène. La raison en est avant tout le décalage entre les langues écrite et parlée. Dans la langue écrite (ou standard), le défini s'exprime généralement par l'utilisation de la forme définie de l'adjectif. Les substantifs sont considérés comme définis *a priori*, les verbes connaissent des formes conjuguées ou définies et indéfinies, etc. C'est précisément l'utilisation de la forme définie de l'adjectif qui pose le plus de difficultés aux locuteurs natifs du slovène.

La désinence *-i* a pour ainsi dire disparu de la majorité des échanges familiers et a été remplacée par l'article *ta*. En raison de l'absence de relation entre les discours parlé et écrit, les écarts sont fréquents dans la langue écrite [...]. En outre, la différence entre les langues parlée, où le défini est codifié à l'aide d'un article, et écrite est si grande que les locuteurs natifs hésitent quant à la forme à employer s'ils n'ont pas appris ce point durant leur scolarité (Schlamberger Brezar 2004 : 42-43).

Si nous concentrons notre attention sur l'analyse du défini dans la langue parlée spontanée (par exemple celle des locuteurs de Ljubljana), nous remarquons que celle-ci est saturée d'actualisateurs généraux, c'est-à-dire d'articles définis et indéfinis. L'analyse précise des textes les plus divers montre ainsi que la langue parlée spontanée (des locuteurs de Ljubljana) exprime le plus souvent le défini et l'indéfini à l'aide d'actualisateurs généraux, ce qui, d'une certaine manière, la rapproche de nombreuses langues romanes et germaniques où l'usage des articles défini et indéfini est codifié dans le système même de la langue, mais l'éloigne de la langue écrite standardisée. De ce fait, bien souvent les locuteurs natifs du slovène hésitent au moment de choisir la forme de l'adjectif (définie ou indéfinie) qu'ils doivent utiliser (s'ils n'ont pas bien appris ce point auparavant).<sup>1</sup>

### 1. LES ACTUALISATEURS TEXTUELS

Les articles défini et indéfini appartiennent à la catégorie des actualisateurs textuels dont la fonction principale consiste à transformer la valeur lexicale du lexème qu'ils précèdent en référence concrète (signification) ou en orientation vers le caractère concret/non concret ou général de la référence présentée. Ainsi, à l'intérieur même du texte, l'indéfini et le défini sont identifiés précisément en raison de leur opposition soulignée par les différences entre les actualisateurs au moyen desquels ils s'expriment (voir : Vidovič Muha 1996: 115-116).

\* Adresse de l'auteur : Filozofska fakulteta, Oddelek za slovenistiko, Aškerčeva 2, 1000 Ljubljana, Slovénie. Mél: mojca.rogac@guest.arnes.si

<sup>1</sup> « /N/ous pourrions nous demander si le codage de la catégorie du défini dans la langue écrite des locuteurs natifs du slovène, tel qu'il est décrit dans les grammaires et manuels de slovène, a encore valeur de système » (Schlamberger Brezar 2004: 43).

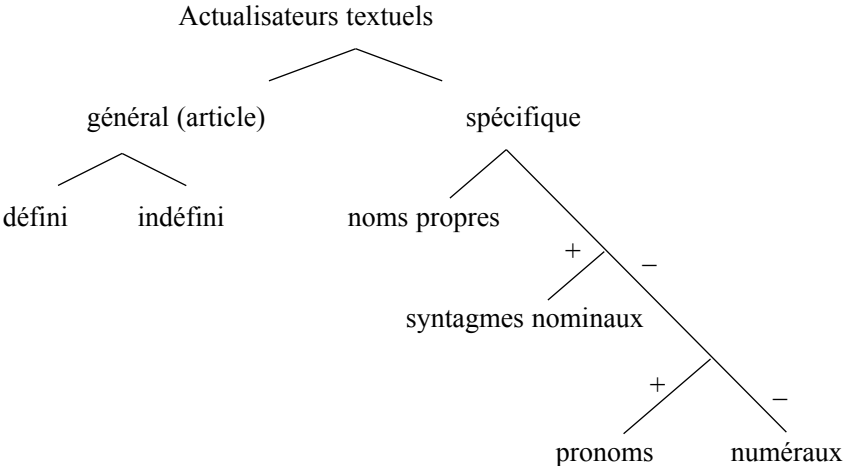
Il s'agit de distinguer des éléments provenant de la relation entre l'indéfinitude propre au système linguistique et la définitude propre à la parole, des éléments provenant donc d'une relation intralinguistique, du « langage ». Pour dénommer cette relation, nous pouvons avoir recours à l'hypéronyme *définitude linguistique*. Comme nous l'avons déjà dit, cette relation se base sur l'indéfinitude propre au système linguistique – sémantisme lexical (virtualité) et la définitude propre à la parole. Dans des textes concrets, cette définitude peut se réaliser soit comme définitude soit comme indéfinitude, les deux étant bien entendu intratextuelles, susceptibles d'être déterminées au niveau de la définitude de la parole sur la base de leur opposition mutuelle (définitude textuelle / indéfinitude textuelle) et discriminables grâce à la différence de leurs supports – actualisateurs. (Vidovič Muha 1996: 115-116)

La principale fonction des actualisateurs étudiés est donc l'expression de la catégorie sémantique de mise en situation spatio-temporelle.

Du point de vue du sens, cette information est définie de trois manières différentes : par son contenu ou sa qualité (« compréhension »), par son extension dans l'espace, qui lui fournit des éléments de quantifications, et par son rapport spatial (et temporel) envers moi, mes discours et mes interlocuteurs (Mikuš 59/60: 11).

La fonction des actualisateurs peut ainsi être assumée par différents moyens linguistiques. Ils se départagent en deux principaux groupes : les actualisateurs généraux et les actualisateurs spécifiques. Sont classés dans le groupe des actualisateurs généraux les articles définis et indéfinis et dans celui des actualisateurs spécifiques les noms propres (qui ne sont dotés de leur valeur informative qu'une fois utilisés dans le texte), les syntagmes nominaux (dont sont exclues les locutions phraséologiques constitutives du dictionnaire), les pronoms et les numéraux.

Schémas : Les actualisateurs textuels<sup>2</sup>



<sup>2</sup> La classification proposée par le schéma est reprise de : Vidovič-Muha (1996: 116-117).

La différence fondamentale entre les actualisateurs textuels généraux et spécifiques réside dans le fait que les premiers sont porteurs des trois éléments actualisateurs (le contenu, la quantité et l'espace/temps). Cela signifie que l'article est en lui-même asémantique et peut ainsi à la fois transmettre et dévoiler des éléments actualisateurs qui caractérisent, quantifient et localisent. À l'opposé des actualisateurs généraux, les actualisateurs spécifiques soulignent explicitement un seul des éléments actualisateurs tandis qu'ils établissent avec les deux autres une relation implicite.

### 1.1. *En et ta*

Dans la région centrale de la Slovénie, ce sont surtout *en* et *ta* qui assument la fonction d'actualisateurs textuels généraux (dans le discours spontané).

#### Exemples :

Ja, ne, to je tko kot, a veš, sta se vidla čez akvarij, a veš, un film Romeo in Julija, a veš, v **tej ta zadni verziji**, k je Leonardo Dicaprio, no tam je ena scena, k se vidta čez akvarij (B12/45).

Dans cet exemple, l'adjectif démonstratif et l'article défini sont utilisés conjointement. Le premier est déclinable (*v tej verziji*) et entre en relation avec le noyau substantival. Du point de vue fonctionnel, il se réalise avant tout au niveau de la référence à la réalité extralinguistique. Le second *ta* (article) n'entre pas en relation avec le noyau substantival ; il est le signe que l'adjectif qu'il précède est déterminé (*ta zadni*).

Tam je vse tko mrtvo, žive duše ni, vsi so že šli, turisti, več al manj zgleda, da domov, ne midva z Nastjo sva rekla, sva mogla še tle za Samota **ta star televizor** dol prinest, pol pa, ne vem, kva je še naroču (B11/37).

Ici aussi, la fonction du lexème « *ta* » est très nette. C'est le signe du caractère défini de l'adjectif qu'il précède.

Pol je, pol je ob pol devetih je bil Boštjan že tam. Pride **en**, a je to stažist, al kaj jaz vem kaj, **en mlad dohtar**, k še niti ni dohtar, mislim, da je stažist, ne, pa tud če ne, no: »Ja, bomo kar mehur prepiknil, ne« (B1/5).

In pol reče: »Ja ja, zdaj bomo pa mi še nekam šli. No, kam naj gremo?«, ne, a veš, pol smo šli še **v en kafič**, tko k ga je Janez poznal in je Janez kar mislu, da bomo tam cel večer zabil (B2/7).

Dans les deux exemples, l'article *en* est le signe de l'indéfini, le signe annonçant que le locuteur ne développera pas le lexème qui suit l'actualisateur (il ne fournira pas de renseignements complémentaires).

Si nous observons tous les exemples cités du point de vue distributif, nous pouvons conclure (peut-être un peu hâtivement) que l'article indéfini est lié du point de vue fonctionnel au substantif (les substantifs expriment l'indéfini à l'aide de l'article

*en/ena/eno*), tandis que l'article défini est lié à l'adjectif (les adjectifs expriment le défini à l'aide de l'article *ta*). Ce que nous venons d'écrire pourrait prouver que les articles défini et indéfini en slovène suivent une évolution indépendante et ne sont en aucun cas un germanisme.

Dans de tels exemples, en réalité nous éprouvons donc le besoin d'utiliser un article défini, et c'est toujours ce que fait le peuple slovène. A-t-il repris cette habitude des Allemands ? Non, car il ne l'utilise pas là où les Allemands l'utilisent, c'est-à-dire devant les substantifs ; il ne l'utilise que dans les cas où les anciens Slovènes l'utilisaient, selon les besoins de leur langue locale (Škrabec 1994: 210).

Si nous partons de cette citation, nous pouvons donc conclure que l'« interruption » de l'emploi des articles défini et indéfini chez Ravnikar<sup>3</sup> est non fondée et que, d'une certaine manière, jusqu'à nos jours elle est à l'origine des hésitations des locuteurs natifs concernant l'usage correct/approprié de la forme définie de l'adjectif et l'expression du défini et de l'indéfini en général.

Dans les œuvres de Ravnikar (*Sgodbe svetiga Pisma sa Mlade Ljudi*, 1815), les articles défini et indéfini disparaissent de la langue standard écrite [...]. On doit sans aucun doute y voir dans une certaine mesure l'influence de Kopitar. Entretemps est parue sa grammaire qui, sur les plans phonétique, morphologique, syntaxique et lexical, a orienté la codification de la langue slovène écrite vers un purisme rigoureux dû au rapport à l'allemand, tout en accroissant l'attention portée aux dialectes slovènes, au vieux slave et aux autres langues slaves (Orožen 1972: 112).

### **1.1.1. L'actualisateur *ta***

En principe, le lexème *ta* assume le rôle d'article défini uniquement quand il est mis en relation avec un adjectif. Dans ce cas, la forme est invariable (*ta/ta/ta*). Dans les textes (retranscrivant le discours spontané), on remarque souvent que l'actualisateur *ta* précède aussi le substantif employé seul (sans adjectif épithète).

Exemple :

M: Skratka eno štorijo ti bom povedou o enmu profesorju fizike z oxfor- oxfordske univerze, k ga dobr puznam in k je tak ful štoser, in predlanskem je mel pač v enem letniku enga študenta, s katerim se res ni zastopu, ne, in tud un ga ni maru, tko da je blo čist reci-pročno, ta ta ... in enkrat je bil spet en konflikt, v polni predavalnci je blo, ne, in kar naenkrat **ta profesor** udpre predal na svojem katedru, potegne ven pištolo in ga ustrelji,

---

<sup>3</sup> Matevž Ravnikar (20. 9. 1776 – 20. 11. 1845) fut le fondateur de la chaire de langue slovène à la Faculté de théologie de l'Université de Ljubljana. Il était également professeur de philosophie et enseignait la Bible. Son œuvre la plus connue est l'ouvrage intitulé *Sgodbe svetiga Pisma sa Mlade Ljudi* (*Histoires de la Bible pour la jeunesse*).



### tega študenta, in ta študent ...

S: A to je res, al kaj?

M: To je res, to je čist avtentična štorija, **ta študent** pade, raide mort, kot bi rekli v francoščini, in pač totalna panika v predavalnici, saj slišal »Oh, God,« a veš, pa »ka pa je zdaj to, saj se mu je čist strgal,« ne, in begat sem pa tja, da ne bi še kerga pihnu, ne, no, in čez kakšno minuto se pa **ta študent** lepo pobere, otrese prah s sebe in se usede nazaj kot da ni nič, in sta se pač očitno, klub temu da se po- čist zares nista marala, sta se zmenila za štos, k sta pač vedla, da vsi vejo, da se ne marata, in da bo to zlo avtentična šala, da bojo vsi verjel, in so dejansko cela predavalnica je padla na finto, ne, in panika, in vse, pa nič, ne, po so pa naprej mel predavanje. Ampak pištola je bla čist *taprava*, sam je bil slep naboj, logično. (B33/102)

Dans cet exemple, les lexèmes en in *ta* sont utilisés conjointement. Ils précèdent tous les deux des substantifs qui ne sont qualifiés par aucun adjectif ou autre déterminant. Il est intéressant de constater que les articles *en* et *ta* sont reliés avant tout aux substantifs se rapportant aux personnages principaux du récit (ou le noyau du récit). Ce sont le *professeur* (*profesor*) et l'*étudiant* (*študent*). Lors de leur première occurrence, ils sont précédés de l'actualisateur *en* (*bom povedou o enmu profesorju fizike z oxfordorske univerze [...] in predlanskem je mel pač v enem letniku enga študenta*). Par la suite, ils sont systématiquement précédé par l'actualisateur *ta* (*in kar naenkrat ta profesor udpre predal na svojem katedru, potegne ven pištolo in ga ustrelji, tega študenta, in ta študent /.../ ta študent pade*).

Le locuteur a utilisé l'actualisateur indéfini également devant certains autres substantifs (*eno štorijo, v enem letniku, en konflikt*). Tous les trois sont mentionnés pour la première fois et le locuteur ne les complète pas, n'ajoute aucune précision les concernant (*v enem letniku - inscrit dans une année d'études* → peu importe laquelle). Tous les trois forment le noyau de la réalité constitutive de la proposition du texte ou du monde textuel.

En dehors des substantifs susmentionnés, précédés d'articles, nous rencontrons dans les textes onze autres substantifs qui ne sont précédés d'aucun article ou actualisateur : *univerza, predavalnica, predal, kateder, pištola, prah, štos/šala, finta, panika, predavanje, naboj*. Traduits en français, ces substantifs sont précédés d'un article : l'université, la salle de cours, le tiroir, la chaire, un pistolet, la poussière, la panique, le cours, une balle. L'absence d'article devant les substantifs énumérés précédemment nous dévoile partiellement les lois régissant l'utilisation des articles défini et indéfini dans la langue parlée spontanée des locuteurs de Ljubljana tels qu'on les rencontre dans l'exemple analysé.

Le substantif *univerza* – *université* est défini déjà à l'aide de l'épithète *oxfordska* – *d'Oxford*. De ce fait, il est impossible de le faire précéder d'un article indéfini ; dans le cas contraire, on pourrait comprendre qu'il existe plusieurs universités d'Oxford. L'utilisation des mots *predavalnica, pištola* et *naboj* dévoile complètement le caractère originellement défini des substantifs slovènes. L'utilisation de l'article indéfini ne serait possible que si l'adjectif épithète était mentionné (par exemple, dans une structure du type : *v eni polni predavalnici, potegne ven eno staro pištolo*). Lors de la seconde mention du mot *predavalnica* (*totalna panika v predavalnici*), le substantif

est déjà défini par le contexte textuel. Il a déjà été mentionné auparavant, mais, comme il ne participe pas du noyau de la proposition du texte, il serait inattendu de le faire précéder d'un article défini. Le substantif *naboj* est déjà déterminé par une épithète (*slep naboj*). Ici aussi, il est impossible de le faire précéder d'un article indéfini, parce qu'il n'y a eu qu'une seule balle - *naboj* - et que le mot est défini par le contexte textuel par la mention antérieure du pistolet - *pištola*. Et pourtant, malgré cela l'utilisation de l'article défini est impossible (*\*sam je bil taslep naboj*). En effet, même si l'ordre des mots ne le confirme pas complètement, *naboj* assume la fonction de sujet et l'adjectif *slep* celle de prédicat. Le mot *kateder* est défini au niveau textuel par l'adjectif possessif qui le précède. L'usage de l'article défini serait redondant, inattendu : comme il ne fait pas partie du noyau de la proposition du texte, le substantif se trouverait particulièrement mis en évidence par la présence de l'article. Le mot *predal* dévoile, lui aussi, le caractère originellement défini des substantifs. L'article indéfini dans le texte pourrait nous faire comprendre que la *chaire* - *kateder* - comporte plusieurs tiroirs. L'article défini serait également inattendu, car *predal* ne représente pas le noyau de la proposition du texte.

Ayant une fonction englobante, les substantifs *štos* et *finta* ne pourraient être précédés que d'un article défini, et pourtant le locuteur l'a omis. Par ailleurs, la locution *pasti na finto* est figée et, de ce fait, ne prévoit pas l'utilisation d'un actualisateur.

Si nous résumons tout ce qui vient d'être énoncé, nous pouvons formuler les affirmations suivantes :

- L'article défini peut également être utilisé devant un substantif ou une locution substantivale.
- L'utilisation des articles défini et indéfini devant un substantif ou une locution substantivale et souvent liée aux personnages principaux du récit, aux personnes constitutives du noyau de la proposition du texte, c'est-à-dire les substantifs ou locutions substantivales présentant/caractérisant les personnes (personnages), les choses, les concepts sur lesquels se concentre le récit.
- Les personnages secondaires (également les noms d'objets et de concepts) sont généralement utilisés sans article (défini ou indéfini), ce qui confirme que les substantifs slovènes sont par définition définis.
- L'utilisation des articles défini et indéfini diffère assez de leur utilisation en français et en allemand. L'utilisation de l'article défini devant un substantif ou une locution substantivale en slovène est plus un procédé stylistique qu'un élément morphologique,<sup>4</sup> puisque les formes peuvent être utilisées avec ou sans article.
- Souvent l'utilisation de l'article défini devant un substantif ou une locution substantivale se recoupe partiellement avec d'autres fonctions métatextuelles ou discursives (par exemple, en tant que moyens anaphoriques ou cataphoriques, en tant que signifiant du rapport au contenu, etc.).

---

<sup>4</sup> Voir : Rudolf Kolarič, *Določni in nedoločni spolnik v slovenščini*, 1961/62.

- L'utilisation de l'article défini devant un adjectif (par exemple, *v tej ta zadni, ta star*) est morphosyntaxique : dans la langue parlée (des locuteurs de Ljubljana), le caractère défini ou indéfini de l'adjectif est encodé presque exclusivement à l'aide de l'article. Même lorsque l'adjectif précède le substantif et forme avec lui une locution substantivale (*v tej ta zadni verziji, ta star televizor*), l'article défini est relié à l'adjectif pour exprimer son caractère défini (il s'agit d'un phénomène morphosyntaxique) et, indirectement aussi, bien sûr, le caractère défini de la locution adjectivale au niveau textuel.

## 1.2 Autres actualisateurs textuels

En dehors des articles *en* et *ta*, on distingue dans le discours parlé spontané de nombreux autres actualisateurs que l'on pourrait appeler (de même que les deux articles généraux) des éléments métatextuels ou signifiants discursifs. Ce sont, par exemple : (*en*) *tak*, *tist(i)*, *un*, *nek*, *tale*. Ils remplissent les fonctions les plus diverses, pouvant être des moyens anaphoriques ou cataphoriques, des déictiques, des marqueurs caractérisant le rapport au contenu, des éléments limitant la validité, des moyens interrompant l'axe linéaire, des déictiques secondaires, des éléments caractérisant l'univers commun, des éléments permettant l'expression du caractère défini ou indéfini au niveau textuel, etc.

Tableau: Les fonctions textuelles et interpersonnelles de quelques actualisateurs textuels

Actualisateurs textuels	Fonctions textuelles et interpersonnelles							
	Fonction anaphorique	Fonction cataphorique	Fonction déictique	Fonction de caractérisation du rapport au contenu	Fonction limitant la validité	Fonction interrompant l'axe linéaire	Fonction de déictique secondaire.	Fonction de caractérisation de l'univers commun
<i>(en) tak</i>	*	*	*	*	*		*	
<i>tisti</i>	*	*	*	*	*	*	*	
<i>un</i>	*	*	*	*		*	*	*
<i>nek</i>				*	*			*
<i>tale</i>	*	*	*	*		*		*

La plupart des actualisateurs étudiés assument plusieurs fonctions à la fois. Par exemple, l'actualisateur *tist(i)* peut agir sur le plan fonctionnel comme un moyen cataphorique et anaphorique. Il peut également caractériser le rapport du locuteur au contenu (le locuteur dévoile son opinion subjective concernant ce qui est dit), il peut limiter la validité, ce qui permet locuteur d'exprimer ses doutes concernant le carac-

tère approprié ou inapproprié du lexème choisi ou bien concernant l'appellation donnée au lexème. Il peut aussi être un procédé interrompant l'axe linéaire ou un déictique secondaire, lorsqu'il transporte *hic et nunc* le monde textuel éloigné dont parle le locuteur. Enfin, il peut également caractériser l'univers commun quand le locuteur exprime ainsi la conviction que le récepteur sait de quoi il parle et qu'il n'est pas nécessaire de lui expliquer le sens du lexème que l'actualisateur précède.

### Exemple:

Recimo, no, recimo, je to taka lušna, k je bil Val majčkn, saj to boste vi tud vidl, takrat k bo Leon mal zrastu, za ene tolk, k je bil že kar ene dve tri leta star, in enkat sedi tko na kahlci, ne, in jaz... to v Martulku, ne, in jaz mu kupim en Kinder jajček, je že takrat norel za temi Kinder jajčki, ne, in on je spet to reku Kindel solplajs, in je skratka ... sedi na kahlci pa prav »: Dida, odvi mi, « ne, jaz pa v **tisti** naglc, mislim, da sem hotu prpravt mu **tist** njegov zrezek, ne, k je stalno meso jedu, ne, rečem »: Joj, Val, ti bom pol, ne, rajši, ne. Zdele nimam časa. « Pa prav »: Ne, kar odvi. « In sem mogu odvit, ne, potem je pa reku »: Ja, « čez neki časa, seveda tam notr je **tist** jajček, ne, čokoladnega itak on ni nikol pojedu, ga je kar zraven odložu, pa tko naprej, ne, nakar je **tist** jajček..., je reku »: Ja, pa tud udpri, ne. « No, po sem pa še tist, sem se udal v usodo, ne, sem reku, zdaj bo dal mir, k bo vidu, kaj je notr, kakšna stvar, ne, kaj za sestav, al, ne vem, ne, in skratka, je po tem no, grem jaz nazaj h **tisti** svoji kuhi, ne, nakar on prav »: Dida, na zdaj mi pa puvej, kako se to sestav, bova sestavla, ne. « Jaz sem reku »: Val, veš kaj, nimam špeglov tuki, ne. « Potem se pa zlo zamisl, pa neki časa je tih, potem pa prav, ko gleda **tist** listek, ne, na kterm je blo pač shema za sestav tale **tisto** igrcu v Kinder jajčku oziroma **tist tist** predmetek, ne, in prav »: Dida, tuki piše: Ne jabiš očal. « (B34/103)

Dans cet exemple, le locuteur utilise à huit reprises l'actualisateur *tisti/a*. Au niveau fonctionnel, les emplois de l'actualisateur étudié diffèrent les uns des autres. Le premier *tisti* (*v tisti naglci*) souligne la signification du mot avec lequel il entre en relation, mais, en même temps, il est le signe de l'éloignement temporel de l'événement (il s'agit d'une action passée) dont parle le locuteur. Le deuxième *tisti* (*tist njegov zrezek*) souligne explicitement la signification du mot auquel il se rapporte. L'adjectif possessif *njegov* a également pour fonction de souligner et de mettre en valeur le substantif. Ensemble, ils expriment donc une action répétitive et évidente, la préparation de l'escalope (*priprava zrezka*). Le troisième *tisti* (*tist jajček*) exprime un fait connu de tous qu'il dévoile et souligne à l'aide de la particule modale *seveda*. Le quatrième *tisti* assume au moins deux fonctions conjointement. Il joue un rôle anaphorique et est ainsi l'élément qui permet de remonter le temps ; en même temps, il est l'élément qui met en valeur le sens du mot avec lequel il entre en relation. Le cinquième *tisti/a* (*h tisti svoji kuhi*) est sur le plan fonctionnel proche du deuxième actualisateur *tisti*. Avec le possessif, il souligne le sens du mot auquel il se rapporte. En même temps, nous pouvons affirmer qu'il joue aussi un rôle anaphorique (1. *sem hotu prpravt mu tist njegov zrezek* → 2. *grem jaz nazaj h tisti svoji kuhi*). En dehors de souligner le sens du mot qu'il précède, le sixième *tisti* (*tist listek*) indique égale-

ment que nous avons affaire à une réalité objective connue de tout le monde ou (selon le locuteur) connue de son interlocuteur. Sur le plan du sens, les septième et huitième *tisti/a* (*tisto igrco*, *tist predmetek*) peuvent être interprétés par la périphrase « *saj veste katero/katerega* ». Sur le plan fonctionnel, ils sont donc proches du troisième actualisateur *tisti* par lequel le locuteur exprime la conviction que son interlocuteur sait de quoi il parle et qu'il est inutile de lui expliquer le sens du lexème précédé par l'actualisateur.

Nous pouvons clore cet examen attentif de l'emploi de l'actualisateur *tist/a* en concluant que l'actualisateur étudié précède, lui aussi, le plus souvent les personnages principaux du récit, c'est-à-dire ceux qui constituent le noyau de la proposition du texte.

## 2. Conclusion

L'expression de la détermination, et plus spécialement de la détermination des adjectifs, cause des difficultés même aux locuteurs natifs avertis de la langue slovène. La désinence *-i* pour exprimer le défini dans la langue parlée spontanée (surtout dans la région centrale de la Slovénie) est remplacée par l'article *ta*. On peut se demander, s'il est vraiment justifié d'affirmer que l'article *ta*, évincé il y a 150 ans, est inacceptable en slovène standard ou soutenu.<sup>5</sup> La problématique étudiée mérite incontestablement de faire l'objet d'un examen plus approfondi fondé tant sur des faits historiques et sociolinguistiques que sur les caractéristiques morphosyntaxiques contemporaines.

Dans les textes retranscrivant le discours direct spontané, l'article *ta* intervient aussi devant les substantifs et locutions substantivales non précédés d'un adjectif épithète ou autre déterminant. L'usage est le plus souvent lié aux éléments du texte qui constituent le noyau de la proposition du texte. Souvent, il se recoupe aussi avec les autres fonctions métatextuelles ou discursives (par exemple, celle de moyen anaphorique ou cataphorique, celle d'élément caractérisant le rapport au contenu ou au récepteur, celui d'élément caractérisant l'univers commun, etc.). En tout cas, nous pouvons conclure que l'utilisation de l'actualisateur *ta* précédant le substantif seul est un moyen stylistique dépendant pour une large part de nombreux facteurs conditionnant l'acte discursif.

## SOURCES

34 documents enregistrés rapportant le discours oral de locuteurs s'exprimant dans le cadre d'une conversation privée. Une fois retranscrits, ces enregistrements comprennent au total 148 427 signes (espaces non comprises). Tous les locuteurs sont originaires de Ljubljana ou de ses environs.

---

<sup>5</sup> Le slovène écrit standard a sans raison rejeté cet article défini. Maintenant que l'on distingue à l'intérieur de la langue standard la langue soutenue et la langue familière, nous pouvons dire que, dans la langue familière, les formes avec *ta* alternent assez librement avec les formes sans *ta*. Pour la langue soutenue, ce *ta*, 150 ans plus tard, n'est pas acceptable (il ne peut apparaître que dans des citations, comme une particularité stylistique dans quelque conte populaire ; Finžgar l'utilise dans ses contes pour souligner le caractère populaire du discours) (Toporišič 1978: 299).

## BIBLIOGRAPHIE

- KOLARIČ, Rudolf (1961/62) « Določni in nedoločni spolnik v slovenščini. » *Jezik in slovstvo* 5/2, 40-44.
- MIKUŠ, Radivoj F. (1959/60) « Prostorni podatak događaja : teorija i govorni izraz. » Radovi, Razdio lingvističko – filološki 1. Zadar: Filozofski Fakultet.
- MIKUŠ, Radivoj F. (1972) « Principes de syntagmatique » Paris: Didier.
- OROŽEN, Martina (1972) « K določnemu členu v slovenščini. » *Slavistična revija* 20/1-4, 105-114.
- SCHLAMBERGER BREZAR, Mojca (2004) « Diskurzivni pristop k problemu določnosti v slovenščini. » *Jezik in slovstvo* 49/5, 35-44.
- SMOLEJ, Mojca (2006) *Vpliv besedilne vrste na uresničitev skladenjskih struktur (primer narativnih besedil v vsakdanjem spontanem govoru)*. Thèse de doctorat. Ljubljana: Filozofska fakulteta.
- ŠKRABEC, Stanislav (1994) « Nekoliko slovenske slovnice za poskušnjo. » In : *Jezikoslovna dela* 2 (XIV. tečaj, 3. zvezek). Nova Gorica: Frančiškanski samostan Kostanjevica, 208-210.
- TOPORIŠIČ, Jože (1978) « Imenska določnost v slovenskem knjižnem jeziku. » *Slavistična revija* 26/3, 287-303.
- VIDOVIČ MUHA, Ada (1996) « Določnost kot besedilna prvina v slovničnem opisu slovenskega jezika (ob Kopitarjevi slovnici). » In : *Jernej Kopitar in njegova doba*, 115-130.

### Résumé

#### QU'EST-CE QUE LE SLOVÈNE ET LE FRANÇAIS ONT EN COMMUN ? LES ARTICLES DÉFINI ET INDÉFINI

On peut distinguer deux principaux groupes d'actualisateurs textuels : les actualisateurs généraux et les actualisateurs spécifiques. Le premier groupe comprend les déterminants définis et indéfinis, tandis que le second est constitué des noms propres, des locutions substantivales, des pronoms et des numéraux.

Le slovène standard exprime le défini, entre autres, grâce à la forme définie de l'adjectif (dans les locutions figées, après les adjectifs démonstratifs et possessifs, quand il s'agit d'évoquer une particularité déjà connue ou évoquée, etc.). De leur côté, les substantifs sont en principe toujours définis. Cependant, si nous analysons l'expression du défini et de l'indéfini dans la langue parlée spontanée (par exemple, des locuteurs de Ljubljana), nous pouvons remarquer que cette dernière se différencie souvent de la langue standard par la présence généralisée des actualisateurs généraux, autrement dit des déterminants définis et indéfinis. L'analyse précise de textes variés montre ainsi que la langue parlée spontanée des locuteurs donnés exprime le plus souvent le défini et l'indéfini à l'aide d'actualisateurs généraux, ce qui, d'une certaine manière, la rapproche des langues romanes et germaniques où l'expression du défini et de l'indéfini est déjà codifiée dans le système de la langue standard.

Povzetek  
KAJ IMATA SLOVENŠČINA IN FRANCOŠČINA SKUPNEGA?  
DOLOČNI IN NEDOLOČNI ČLEN

Besedilni aktualizatorji se delijo na dve večji skupini, na skupino globalnih aktualizatorjev in skupino specialnih aktualizatorjev. V skupini splošnih so določni in nedoločni členi, v skupini posebnih pa so lastna imena, imenske zveze, zaimki in števniki.

Knjižna slovenščina določnost izraža npr. z določno pridevniško obliko (v stalnih zvezah, za vrstnimi kazalnimi in svojilnimi zaimki, ko gre za znano oz. že omenjeno lastnost...), samostalniške besede pa so načeloma tako in tako določne. Če se osredotočimo na analizo določnosti v spontano govorenem jeziku (npr. govorcev Ljubljane), lahko opazimo, da se le ta pogosto razlikuje od knjižnega jezika, saj je v njem prisotnih vse polno globalnih aktualizatorjev oz. določnih in nedoločnih členov. Natančna analiza najrazličnejših besedil tako pokaže, da spontano govoreni jezik (npr. govorcev Ljubljane) določnost oz. nedoločnost največkrat izraža prav z globalnimi aktualizatorji, kar ga na nek način približuje številnim romanskim in germanskim jezikom, kjer je raba določnega in nedoločnega jezika zakodirana v sam sistem knjižnega jezika.





**COMPTEs RENDUS, RÉCENSIONS, NOTES**  
**POROČILA, OCENE, ZAPISI**

**Barbara Sandig. Textstilistik des Deutschen. Berlin/New York: Walter de Gruyter. 2006. xiii + 584 S. ISBN : 9783110188707**

Nach den allgemeinen Einführungsthemen wie Stilbeschreibung und Textstilistik wird das „Textphänomen“ (11) Stil als (sozial bedeutsame) Art der Handlungsdurchführung definiert (9, 17), die aufs Engste mit der Stilkompetenz verbunden ist. Aus der Opposition der Struktur und Funktion leitet die Autorin den Begriff des stilistischen Sinns als eine spezifische Bedeutung einer Textstruktur ab (11) und stellt diesem die Stilwirkung als pragmatische Größe gegenüber, die durch die Interpretation von Handlungsbeteiligten hergestellt wird (24). So wie bei stilistischem Sinn (11) arbeitet sie verschiedene Typen von Stilwirkungen heraus (34).

Im 2. Kapitel („Die generelle Stilstruktur“) geht es um den Stil als Bündel konkurrierender Merkmale, wobei auch Mischungen (58) und internen Strukturen von Merkmalsbündeln (gegenüber Wahlmöglichkeiten und Spielräumen (65)) Aufmerksamkeit geschenkt wird. Das Fortführen von Merkmalen (65) mit stilistischer Qualität sei dabei ein spezieller Typ der Textherstellungshandlung. Zur Stilstruktur gehören auch Stilgestalten, die Diskontinuitäten aufweisen (71) und aus dem Wissen der Rezipienten „mühelos“ ergänzt werden (Frames). Stilistische Gesamtgestalten können aus mehreren relationierten Teilgestalten gebildet werden (76 f.). Analysiert werden drei Varianten: das Nacheinander (Thema-Rhema-Gliederung und Rhythmus), das Miteinander (Parallelismen durch Rekurrenz, Zusammenspiel von Stilmerkmalen verschiedener Beschreibungsebenen (81)) und das Übereinander stilistischer Teilgestalten (Synergie von Teilgestalten aus verschiedenen Beschreibungsebenen).

Das 3. Kapitel „Kontexte: Stil ist relational“ handelt vom Vergleichen als einer zentralen Methode der Stilbeschreibung (85). Interpretationen sollten vor diesem Hintergrund in einer Interaktion durchgeführt werden (wobei besonders bei Nichtmuttersprachlern auch Assoziationen auf kontrastive und konfrontative Aspekte geweckt werden), was entscheidend von verschiedenen Wissensbeständen bei den Beteiligten (87 f.) abhängt. Interne Relationen der Handlungssequenzierung (aller Varianten der grammatischen und semantischen Textkohäsion mit zahlreichen Ausdrucks-Alternativen (85)) müssen dabei gleichzeitig mit textexternen Relationen (113) (die Kommunikationssituation mit allen – auch kulturellen (135) - Parametern) berücksichtigt werden.

„Textstilistische Handlungsmuster“ (4. Kapitel) sind für Sandig „Muster für das Durchführen stilrelevanter Teilhandlungen, für die eine große Bandbreite stilistischer Merkmale zur Verfügung steht“ (147). Das Zusammenspiel von diesen Merkmalen liegt der stilistischen Wahl aus dem systemangelegten Inventar zugrunde. Teilhandlungen

sind miteinander kombinierbar und kommen in verschiedenen Textsorten vor. Das sind Verfahren zum sprachlichen Realisieren von kommunikativen Intentionen.

Handlungsmuster teilt Sandig in verschiedene Untergruppen ein: in allgemeine textstilistische Handlungstypen und Verfahren, in generelle textstilistische Muster und in komplexe stilistische Handlungsmuster (147 ff.). In demselben Kapitel werden noch typisierte Stile anhand von Beispielen und Stilebenen beleuchtet. Unter den allgemeinen werden die „allgemeinsten“ als Textherstellungshandlungen par excellence hervorgehoben, in erster Linie das Fortführen (151) (Wiederholung, Rekurrenz, Phorik). Das Wiederholen ermöglicht auch das Variieren als Grundprinzip der Stilistik (152). Zu den allgemeinen Stilverfahren zählt Sandig vor diesem Hintergrund das Abweichen, das Verdichten, das Mustermischen und das Verfahren „Einheitlich machen und wechseln“. Das Abweichen erfolgt auf allen Beschreibungsebenen (graphisch, lautlich, morphologisch, semantisch, syntaktisch und pragmatisch, textprototypisch (Textmuster-mischungen), typographisch). Unerwartete Wendung, Verletzung der Erwartungsnorm und Schattierung der Interaktionsmodalität sind stilistische Abweichungen im engeren Sinne des Wortes. Auch die behandelten Handlungen des Verdichtens sind sehr variationsreich und weisen eine ganze Palette verschiedener Varianten quer durch das Sprachsystem auf, die jedoch alle miteinander korrelieren und synergetisch wirken. Die Mustermischung ist ein Gradmesser der Invention der Sprachproduzenten und macht sich vor allem in Diskursen bemerkbar, die auf besondere Effekte bei den Rezipienten hinzielen (Werbesprache).

Zu den eher formalen textstilistischen Mustern gehören generelle textstilistische Muster, die „variabel einsetzbar“ sind (175). Hiermit meint Sandig vor allem verschiedene strukturelle Verfahren, mit Hilfe derer auch rhythmische (180) und kommunikativ-informative Akzente gesetzt werden können (funktionale Satzperspektive) (178). Kontrastieren und Pointieren werden dadurch ermöglicht. Dabei spielen auch die Erkenntnisse der Kognitionspsychologie (Epistemologie, 197) und der Natürlichkeitstheorie (Ikonizität, 196) eine wichtige Rolle (186). Diese Erkenntnisse schlagen sich schon wieder auf allen sprachlichen Beschreibungsebenen nieder. In einschlägigen Textanalysen wird darauf hingewiesen, dass diese Phänomene nicht für alle Textsorten von gleicher Relevanz sind (198). Hervorgehoben wird weiterhin das Kontrastieren als ein Verfahren mit langer stilistischer Tradition (206). Nicht nur auf der Mikroebene können z.B. die Stilfiguren Antithese und Chiasmus stilistisch wirksam eingesetzt werden. Das kann auch ein „Aufbauprinzip von Absätzen und Texten sein“ (206). Implizite Vertextungsmuster verstecken sich im Dialogisieren (212) und Referieren (216). Das Hervorheben, die Emphase (235), das Erzeugen der Spannung (243) und das Anschaulichmachen (auch durch Beachtung der Ikonisierungen) (247) sowie die Informationsgewichtung sind weitere Unterarten der generellen stilistischen Handlungsmuster, bei denen ebenfalls die gegenseitigen Abhängigkeiten der sprachlich-systematischen Mikroebene und der pragmatischen Makroebene zum Ausdruck kommen.

Komplexe stilistische Handlungsmuster sind solche, in denen „einfachere Muster integriert sein können“ (249), so dass diese in einer modalen Relation als Mittel zum Zweck der übergeordneten Handlung dienen. Dazu gehören nach Sandig z.B. Verständlich machen, Bewerten, Emotionalisieren (249), aber auch Perspektivieren (260) (Deixis) (263), Themenentfaltung und Thema-Rhema-Gliederung. (Hier kämen sicherlich auch verschiedene auf Subtexte zurückgehende Kommunikationsebenen mit Mehrfachadressierung der Rezipienten in Frage. Vgl. S. 271.) Das Verständlichmachen als kommunikative Handlung ist aufs Engste mit den Wissensbeständen des Produzenten und vor allem auch des Rezipienten verbunden (vgl. auch Kapitel 5).

Typisierte Stile werden an Hand von Analysen des Bibelstils, des Parlando, des aggressiven Stils (Sprachglosse) und von Jargons präsentiert (279).

Im Unterkapitel 4.6 werden Stilebenen unter einem etwas anderen Gesichtspunkt als Träger eines globaleren Sinnpotenzials definiert und an verschiedenen Textanalysen in dieser Funktion beleuchtet. Das „Neutrale“ (291) wird als eine Art unmarkierter default-Wert antizipiert, von dem dann Abweichungen nach oben und unten als überneutrale und unterneutrale Kombinationen von Merkmalsbündeln aufgefasst und besprochen werden. An einigen Musteranalysen werden anschließend einige wesentliche Merkmale des überneutralen (Feierlichkeit und Pathos, positive Selbstdarstellung) und des unterneutralen (mündliche Alltagsrede) Registers herausgearbeitet (301). Im letzteren Fall wird auf die Entwicklungstendenz aufmerksam gemacht, dass etwa umgangssprachliches Register mit allen verfügbaren Ausdrucksmöglichkeiten auch in den schriftlichen Kode eindringt (300 ff.) und dabei im Zusammenspiel mit den beiden anderen Ebenen in der Verwendung verschiedene kommunikative Funktionen ausübt (Emotionalisierung, Einstellungsausdruck, Lesernähe usw.).

Das umfangreichste Kapitel 5 (179 Seiten) untersucht die Vielfalt stilistischer Phänomene in Texten als Mittel sprachlichen Handelns (307). Ausgegangen wird von der hermeneutischen Auffassung (308), dass Texte „Halbfabrikate“ sind, die „erst vom Rezipienten zu einem kohärenten Gebilde durchwoben werden (müssen)“ (308). Die Autorin unterscheidet folgende Textmerkmale: Unikalität, Funktionalität, Thema, Kohäsion, Kohärenz, Situationalität und Materialität, Kulturalität und Historizität. Diese Merkmale überlappen sich weitestgehend mit den Textualitätskriterien bei de Beaugrande/Dressler (1981), wenn man berücksichtigt, dass die Intertextualität in 5.4.2.3 und an zwei anderen Stellen behandelt wird (die referentielle im Kapitel 3.1.3.2, und die typologische im Kapitel 6), dass Funktionalität mit der Intentionalität, Thema mit der Informativität gleichzusetzen sind und die Materialität eigentlich ein Aspekt der Kohäsion ist.

Die Textualitätskriterien von de Beaugrande/Dressler (ebd.) sind jedoch um die Unikalität sowie Kulturalität und Historizität erweitert, die Ulla Fix schon 2001 postuliert hat. (Fix/Poethe/Yos 2001)

Bei der Unikalität handelt es sich in erster Linie um die Relation Textmuster und konkrete Textrealisierung – inwiefern hat man sich bei der Textproduktion an den Kanon zu halten, wie viel individuellen Freiraum wird einem zugelassen. Texte seien

letztendlich „einmalig vorkommende Handlungen“, als solche müssen sie auch ein individuelles Gepräge erhalten. Eine ganze Bandbreite fließender Übergänge wird herausgearbeitet (z.B. Unikalitätsskala, 313).

Die Funktion ist nach Sandig das zentrale Merkmal von Text. Der Autorin geht es hier nicht um die „textuellen Grundfunktionen“, sondern um die „spezifischen Textfunktionen“ (318), die auch den jeweiligen „sozialen Sinn“ (318) eines Textmusters antizipieren. An konkreten Textsortenbelegen werden verschiedene Hinweise für das Erkennen der Textfunktion (explizite Kategorisierungen, ausgeprägt wahrnehmbare Textgestalten, interne Textstrukturierung, Art des Themas, Textträger und/oder Situationstyp) systematisch veranschaulicht. So wird auch die Problematik undeutlicher und individueller Textfunktionen berührt (331 f.), wo verstärkte Interpretationsleistung seitens der Rezipierenden nötig ist. Eng aufeinander bezogen sind auch das Thema und die Textfunktion (335), und die Themenformulierung weist eine Reihe verschiedener Varianten auf (338 ff.). (Vgl. auch zur Relation von Hauptthema – Unterthema – Nebenthema, 348.)

Der Terminus Textmerkmale scheint geeigneter zu sein als Textualitätskriterien. Diese Textualitätscharakteristika sind in Texten so unregelmäßig vertreten, dass man sie nicht als Kriterien, sondern höchstens als Merkmale bezeichnen kann.

Interessant ist in diesem Zusammenhang die Präsentierung, mit welchen syntaktischen Mitteln Hauptinformationen und untergeordnete Nebeninformationen indiziert werden (350). Damit hängen auch die Themenklassifizierung (vordergründige und hintergründige Themen) und die Prinzipien zusammen, wonach diese Themen und Subthemen miteinander verwoben sind und im Text entfaltet werden (351 f., 357). Diese Entfaltungsmuster sind zugleich stilistische Handlungsmuster. Thematische Irreführungen zeigen sich vor diesem Hintergrund als stilistische Möglichkeiten der Variation, auch in Kombination mit Bildelementen im Text (354). Irreführungen sind aber neben den vier bekannten Grundformen thematischer Entfaltung (nach Brinker (2001) deskriptiv, narrativ, explikativ, argumentativ) im Prinzip einer der weiteren Typen der Themenentfaltungen (zusammen mit dialogischen Themenentfaltungen, unernten Themenentfaltungen, der Themenentfaltung mit überraschender Wendung oder mit einer Pointe u.d.m.) (357). Zu hinterfragen wäre, in welcher Relation dazu etwa die thematischen Progressionstypen der Funktionalen Satzperspektive (einfache lineare Progression, durchgehendes Thema, abgeleitetes Thema, gespaltenes Rhema, thematischer Sprung) stehen (vgl. Eroms 1986).

Sandig macht (berechtigt) einen deutlichen Unterschied zw. Kohäsion und Kohärenz. Die formale und die funktionale Seite der beiden Merkmalsprägungen ermöglichen und bedingen im Regelfall einander (363). Trotz sog. Kohärenzjokern als Kohäsionshinweisen könne es nullkohärente Texte geben (407).

Kohäsionsmittel für die Herstellung der Kontinuität und Diskontinuität werden aufgeführt; unterschieden wird zwischen Junktion, Asyndese, Phorik (Anaphorik, Kataphorik), Rekurrenz (367), Fortführen (377 u. passim). Bei der ausdrucksseitigen Rekurrenz geht die Autorin sogar auf die Elemente der Wortfamilie ein, die dank

eines gemeinsamen Grundmorphems kohäsive Relationen ermöglichen (368; vgl. a. 371). Ein kohäsives Mittel sieht sie auch im Rhythmus, der u.a. durch Reime hergestellt wird (369). Zum besonderen Rhythmus tragen auch Nebenprädikationen bei (375, 380), die im Prinzip indirekte Rhematisierungen (Verdichtungen bei Eroms (ebd.)) sind. Parallelismen, insbesondere Chiasmen, wirken ebenfalls kohäsiv (370). Textdeixis wirkt sich prototypisch in Fachtexten aus (378). Illustrativ ist dabei die zusammenfassende Tabelle, die Grade der Deutlichkeit anzeigt, die durch verschiedene kohäsive Mittel bzw. durch deren Abwesenheit erzielt wird (386). Dem Wissen in seinen verschiedenen Existenzweisen (auch Frames und den z.T. darauf beruhenden pragmatischen Präsuppositionen – 389, 397 und passim) wird die maßgebliche Rolle bei der Herstellung des Sinns eines Textes zugeschrieben, und das ist aufs Engste mit der Themenformulierung verbunden (388). Die implizite Kohärenz beruht nicht selten auf den Varianten der globalen Textstrukturierung (391). Wie die textgrammatischen Verfahren mit Textstil verbunden sind, darauf weist die Autorin u.a. bei der Erwähnung von spezielleren Kohärenzbildungshinweisen (399) hin: Halter, Klammer, Scharnier, Analogisieren. Ein Kapitel für sich ist die Kohärenz von Sprache und Bild. Das wird auch an einigen Werbetexten thematisiert und veranschaulicht (s. auch 452, 460, 168, 460).

Die Situationalität hebt die Wichtigkeit des Situationswissens hervor, sie wird definitorisch als Faktor der Textualität von de Beaugrande/Dressler abgeleitet (1981, 413). Situationstyp und konkrete Situation werden unterschieden, die Rolle des Textträgers als Material, auf welches die Textoberfläche aufgeprägt ist, wird hervorgehoben (413). Eingegangen wird auch auf Typen der Präsuppositionen (416) und auf Typen der Pragmatik in einer kommunikativen Situation (415 ff.). Die Begriffe Kanal, Textträger und Medien werden voneinander abgegrenzt (427 ff.). Auch die haptische Qualität von Textträgern (z.B. Qualität von Papier) spielt eine Rolle (434). Im Lichte der Materialität als Merkmalsausprägung von Texten wird auch Text als begrenzte Einheit unter die Lupe genommen (436); die Strukturen der konkreten Poesie und der visuellen Textgestalt überhaupt (und der darin berücksichtigten und angewendeten Ikonizität) werden beleuchtet (442 f., 450, 459).

Aus der Rezeptionsperspektive gesehen, sind auch die Textualitätsfaktoren und Parameter von Bedeutung, die Sandig unter dem Sammelbegriff Textsequenzierung zusammenfasst (473): Textmerkmale, auf die ein Rezipient bei der Rezeption zunächst stößt und die sich dann nach und nach als relevante Textkonstituenten herausstellen: Typographie, Themenformulierung und –entfaltung (natürlich oder artifiziell?), Anfang-Mitte-Ende-Struktur, Ikonizität u.d.m. Das sind Anhaltspunkte für das Erkennen der Textfunktion (473). Als ein wichtiges Textmerkmal wird in Anlehnung an Fix/Poethe/Yos (ebd.) abschließend die Kulturalität (481, 16) herangezogen.

An zwei Textmustern (Glosse und Heiratsannonce) wird in Kapitel 6 die komplexe Relation zwischen Textmuster und Stil dargestellt. Einleitend wird auf die wichtige Korrelation zwischen der Wahl der Stilmittel und den durch Konventionen geforderten Restriktionen aufmerksam gemacht (487). Dem Zugang bei der Textmusterbeschreibung

und den Beziehungen zwischen Handlungstyp, Textsorte und Erwartungen, die damit verbunden sind, wird Rechnung getragen.

Die „Textstilistik des Deutschen“ aus der Feder von Barbara Sandig ist eine pragmatische Stilistik im weitesten Sinne des Wortes. Obwohl sie vor fünf Jahren erschienen ist, hat sie nichts an ihrer Aktualität verloren. Im Gegenteil, sie wird im Bereich der Fachdisziplin Stilistik wohl einen Ehrenplatz einnehmen, wie etwa im Bereich der Semantik „Deutsche Satzsemantik“ von Peter von Polenz (1988) oder für die Textlinguistik die „Einführung in die Textlinguistik“ von Robert-Alain de Beaugrande und Wolfgang Ulrich Dressler (1981).

Diese Stilistik ist, um einen (anders referentiell besetzten) Terminus aus ihrem Vorwort zu verwenden, ein „Eldorado“ für alle GermanistInnen, ja LinguistInnen allgemein, die die auf zahlreiche repräsentativ ausgewählte Textbeispiele angewendeten theoretischen Prämissen der Stilistik hinterfragen wollen. Den umfangreichen Band (584 Seiten) kann man wegen der Fülle und Dichte der enthaltenen Informationen nicht schnell lesen. Darüber hinaus erfordert das Studium der Theorie und der analysierten Textbelege viel Aufwand, um aufgrund der bisweilen disloziert behandelten Phänomene ein komplexes Bild von der konkreten Problematik zu gewinnen. Viele Querverweise symbolisieren, wie eng stilistische Phänomene miteinander verwoben sind und nur aus system-analytischen Gründen an verschiedenen Stellen im Buch eingeordnet werden müssen. Die Problematik wird konsequent angegangen, die Methodik ist alles andere als langweilig. Man findet darin eigentlich alle linguistischen Themen, die direkt oder indirekt mit Stilistik verbunden sind. Beeindruckende Textanalysen sensibilisieren für stilistische Fragen und erweitern den linguistischen Horizont. Zahlreiche Abbildungen sind zusammen mit guten Übersichtstabellen und –skalen sowie einer Fundgrube an Textbelegen ein anschaulicher Apparat, der die theoretischen Ausführungen zugänglicher macht.

Für eine Neuauflage könnte man sich eventuell ein etwas ausführlicher ausgearbeitetes Sachregister wünschen. Aber das ist schon eine Konzept- oder gar Geschmacksache. Schade, dass das Buch nicht fester gebunden ist. Da man es wegen seines enzyklopädischen Charakters immer wieder in die Hand nimmt, werden aus dem Textträger bald fliegende Blätter.

Stojan Bračič  
Universität Ljubljana

**Toshiko Yamaguchi. *Japanese Linguistics: An Introduction*. London/New York: Continuum. 2007. XVIII + 220 S. ISBN: 9780826487896**

Das hier zur Besprechung vorliegende Werk unterscheidet sich in vielerlei Hinsicht von Büchern ähnlicher Art, die Japanisch zum Gegenstand haben. Der Leitgedanke scheint zu sein, dass der Erwerb des Japanischen als Zweitsprache oder Fremdsprache durch die richtige Anwendung wissenschaftlicher Forschungsergebnisse erleichtert werden kann. Diese werden für den Lernenden aber unbewusst eingesetzt und nicht, wie man erwarten würde, durch vorheriges Definieren erklärt und anschließend angewendet. Um diese Technik zu erreichen werden die Erscheinungen, von denen die Rede ist, anhand von Beispielen aus realen, authentischen Texten exemplifiziert. So wird man sich dessen bewusst, dass es sich um Erscheinungen handelt, die im realen, täglichen Sprachgebrauch vorkommen. Die Fachbegriffe, die notwendig sind, werden in dem jeweiligen Kontext anhand solcher realer, authentischer Beispiele eingeführt, aber nicht wie es üblich ist anhand vorheriger theoretischer Erklärungen.

Als „authentisch“ werden solche Texte verstanden, die nicht für den Zweck der Darstellung verfasst worden sind. Das sind geschriebene Texte, die uns im täglichen Leben begegnen: in Zeitungen, Zeitschriften, Romanen, Gebrauchsanleitungen, Bildbänden und anderen Veröffentlichungen. Diese bilden das Gerüst, in das die grammatischen und fachlichen Erscheinungen integriert sind.

Das Buch besteht aus sieben Kapiteln: Chapter 1 *Speech sounds*; Chapter 2 *Sound Structure*; Chapter 3 *Vocabulary*; Chapter 4 *The Writing System*; Chapter 5 *Word Structure*; Chapter 6 *Word Meaning* und Chapter 7 *Sentence Structure*. Jedes Kapitel enthält Übungen zu den behandelten Themen. Die authentischen Texte werden teilweise übersetzt, aber gelegentlich nur kommentiert, so dass der Lernende gefordert ist, um das besprochene Phänomen zu verstehen. Obwohl einige Abschnitte sicherlich auswendig gelernt werden können, ist Auswendiglernen nicht angestrebt. Die Übungen sollen vielmehr in erster Linie sowohl selbständiges Denken als auch kritisches Lernen fördern und zum Reflektieren über die Sprache anregen. Daher könnte das Buch gleichermaßen als Arbeitsbuch und Lehrwerk charakterisiert werden.

Der Leser entwickelt eine Begeisterung für das Thema. Er entdeckt, dass er die Sprache auf diese Weise nicht nur lernen, sondern auch gleichzeitig verwenden kann. Die Sprache erscheint ihm als ein lebendiges Element, das in seine tägliche Tätigkeit integriert wird. Sie wird somit zu einem natürlichen Bestandteil im täglichen Leben, in welchem sie assimiliert wird. Dieser Assimilationsprozess geht jedoch nicht mühelos vor sich hin, aber der Lernende erlebt jeden Schritt mit Freude und wird nicht den Frust empfinden, der sich häufig bei sprachlichen Lehrwerken bemerkbar macht.

Wenn die Überschriften der einzelnen Kapitel betrachtet werden, scheint das Buch recht traditionell zu sein. Solche oder ähnliche Kapitelüberschriften finden sich in sehr vielen Lehrwerken. Dieser Eindruck täuscht jedoch. Dies ist kein Lehrwerk im traditionellen Sinne. Die inhaltliche Darstellung bereits bekannter Fakten ist vielfach anders und in gewissem Sinne revolutionär. Vermieden werden große Aufzählungen von

Wörtern und Beispielen, die häufig in Lehrwerken zu finden sind. Statt dessen werden die Beispiele langsam eingeführt, in kurzen gut strukturierten Paragraphen erläutert, die der Lernende schnell assimilieren kann und anhand authentischer Beispiele exemplifiziert. Beispiele dieser Strategie lassen sich im Buch fast überall finden. Besonders eindrucksvoll sind in dieser Hinsicht die Kapitel 3, 5, 6 und 7 strukturiert.

Es werden aber nicht nur Daten und Fakten präsentiert. Der Inhalt wird pädagogisch vorbereitet und so dargestellt, dass Lernende das Buch optimal verwenden können. Allerdings werden gewisse Vorkenntnisse vorausgesetzt. Dies gilt insbesondere in zwei Bereichen. Einerseits im Bereich des japanischen Schriftsystems, das hinreichend bekannt sein muss, um die Texte verstehen zu können. Der Lernende muss wissen, wie das Schriftsystem zusammengesetzt ist und wie es funktioniert. Es ist ein Novum, dass hier nur begrenzt mit Transliteration gearbeitet wird. Die Texte werden in der ursprünglichen japanischen Schrift präsentiert. Im übrigen ist zu betonen, dass im Buch streng zwischen Transliteration und phonetisch/phonologischer Transkription unterschieden wird. Dies ist sehr nützlich, um dem Lernenden deutlich zu machen, in welcher grundlegender Hinsicht sich das japanische Schriftsystem von der Lateinschrift unterscheidet.

Andererseits werden in der Morphologie ebenfalls gewisse Vorkenntnisse erwartet. Genaue Erläuterung der Verbmorphologie erfolgt nicht. Erläutert wird die Verwendung der Formen, deren Bildung als vorausgesetzt angenommen wird. Es wird dem aufmerksamen Leser auch aufgefallen sein, dass kein Kapitel die Überschrift *Morphology* trägt. Die morphologischen Formen werden in dem Kontext, in dem sie verwendet werden, eingeführt und die Stelle in ihrer Verwendung erläutert. Die Formen selbst werden als bekannt vorausgesetzt.

Dieses Werk gehört zu einer neuen Generation von Lehrwerken<sup>1</sup>, in denen Daten auf wissenschaftlicher Grundlage und pädagogische Elemente integriert werden. Es ist nicht nur für Japanischlernende ein sehr gutes und nützliches Werk, sondern auch für jeden Linguisten eine wichtige Informationsquelle, die beispielhaft eine Sprache in ihrer Struktur und Anwendung zeigt. Es hat einen Anwendungsbereich weit über den Kreis der Japanischlernenden hinaus. Wissenschaftliche Erkenntnis und pädagogische Eignung sind in diesem Werk hervorragend miteinander verbunden.

Magnús Pétursson  
Universität Hamburg

---

<sup>1</sup> Ein anderes Lehrwerk, das ich auch dieser neuen Generation von Lehrwerken rechnen würde, da es wissenschaftliche Begriffe anhand einfacher Beispiele einführt und so in einfacher Sprache eine sehr bemerkenswerte Genauigkeit erreicht, ist meiner Ansicht nach Heidi Harley: *English Words: A Linguistic Introduction* (Blackwell, Malden (USA) and Oxford (UK) 2006)



**Paul Videsott. *Padania scrittologica. Analisi scrittologiche e scrittometriche di testi in italiano settentrionale antico dalle origini al 1525*. Beihefte zur Zeitschrift für romanische Philologie 343. Tübingen: Max Niemeyer Verlag, 2009. XVII + 624 pp. ISBN: 9783484523432**

Il presente volume, versione rimaneggiata e tradotta della *Habilitationsschrift* dell'autore,<sup>1</sup> mira a “delineare l'evoluzione della lingua scritta volgare non letteraria dell'Italia settentrionale” (7), adoperando un metodo sviluppato, e variamente applicato, soprattutto da Hans Goebel. L'opera si divide in quattro capitoli, preceduti da due prefazioni (1-6)<sup>2</sup> e seguiti dalla “Bibliografia” (423-454), dalle “Cartine coropletiche” (455-620), e da due indici (621-624).

Nei primi paragrafi dell'“Introduzione” (7-267) l'autore espone lo “Scopo del lavoro” (7-9), presenta le “Definizioni” (9-10) dei più importanti termini tecnici utilizzati (*scripta*, scrittologia e scrittometria), si ferma brevemente sullo “Stato attuale della ricerca” (11-13), e giustifica “L'arco cronologico analizzato” (13-17), che va dalle prime attestazioni fino al 1525. Seguono “I fondamenti metodici e metodologici dell'analisi” (17ss), tra cui l'elenco dei “centri scrittori analizzati” (19-23), quello dei 320 criteri linguistici utilizzati per l'analisi (28-48), nonché quello di tutti i 2064 testi contenuti nel corpus (64-237). Tuttavia, per i motivi esposti nei sottoparagrafi successivi, il corpus effettivamente analizzato consiste di 1165 testi, elencati in base ai criteri “centro di scrittura”/“data di redazione dell'originale” (240-256) e “tipo di testo” (257-259). La parte introduttiva si chiude su due quadri sinottici che informano sulla stratificazione cronologico-quantitativa del corpus effettivo (260-264) e sugli archi cronologici coperti dai singoli centri scrittori (265-267).

La seconda sezione del libro è dedicata alle “Analisi scrittologiche” (269-406). Nel paragrafo intitolato “Il metodo usato” (269-279) viene dimostrato come, sulla base di determinati calcoli, si ottenga il valore *Frel* (“Frequenza relativa”), attinente ad un criterio particolare ed ad una determinata parte del corpus, e il valore *Dabs*, cioè “la *Differenza tra la loro* [sc. dei criteri] *occorrenza assoluta e quella prevista teoricamente*” (269). Considerando, ad esempio, tutti i 320 criteri e tutti i 1165 testi, si ottiene un valore *Frel* globale pari al 29,99% (v. 273), il che significa un “grado di ‘non-toscanità’ (inteso come distanza dallo ‘standard’) della *scripta* italiana settentrionale [...] molto marcato, in netto contrasto ad es. con la *scripta* della Normandia”, per la quale, da ricerche analoghe, era risultato “un valore pari al 7,58%” (273). Nel paragrafo successivo (280-289) sono presentati, principalmente con una serie di tabelle, i diversi valori *Frel* che si ottengono a seconda delle diffe-

---

<sup>1</sup> Nella lista delle “Romanistische Habilitationsschriften und Dissertationen” del 2004 pubblicata sul *Romanistisches Jahrbuch* 55, lo studio appare con il titolo tedesco “*Padania scrittologica. Skriptologische und skriptometrische Untersuchungen zu altnorditalienischen Texten von den Anfängen bis 1525*” (15).

<sup>2</sup> Quella alla versione pubblicata e quella alla versione manoscritta.

renti variabili prese in considerazione (‘centro scrittorio’, ‘periodo considerato’<sup>3</sup>, ecc.). Nel terzo paragrafo (290-406) sono discussi, “a titolo esemplificativo” (290), alcuni dei criteri, tra l’altro gli esiti (grafici) della desinenza latina -ĀRIU, del dittongo AU, e del nesso ĀL + consonante.

Il terzo capitolo è dedicato alle “Analisi scrittometriche” (407-418), metodo “per la prima volta” applicato a “una parte dell’area linguistica italiana” (7). Lo scopo di tali analisi consiste nella creazione del “profilo di similarità” tra due centri scrittorios, utilizzando i risultati dell’analisi scrittologica per ottenere, tramite determinate procedure matematiche, una “matrice di similarità” i cui valori numerici indicano la (dis)similarità tra un centro scrittorio scelto come punto di riferimento e gli altri centri. Utilizzando tali procedure diventa possibile condurre ulteriori analisi che permettono l’interpretazione dei dati in chiave isoglossica, nonché “l’impiego di processi gerarchico-agglomeranti” (416).

Il volume chiude con un “Riepilogo dei risultati” (419-422), in cui l’autore sottolinea come il suo “scopo principale” (419) fosse lo stesso di Hans Goebel e Guillaume Schiltz, ossia “über viele Einzeltexte hinweg *globale Strukturen* bzw. *Musterungen* nach Raum und Zeit [zu] untersuchen”<sup>4</sup>. Tra gli aspetti menzionati nella conclusione ci limitiamo a ricordare che, dalle analisi condotte, emerge chiaramente la costante diminuzione dell’assetto non-toscano delle *scriptae* padane, nonostante occorra tuttavia differenziare tale quadro generale: a centri scrittorios molto innovativi come Milano e Venezia, infatti, se ne oppongono, ad esempio, altri più conservatori come Belluno e Udine.

Da quanto appena riferito per sommi capi, è evidente che si tratta di un’opera molto complessa e ricchissima d’informazioni, la lettura della quale, seppur impegnativa, ricompensa però pienamente chi raccoglie la sfida. Se, nel complesso, tale impegnatività è indubbiamente collegata all’oggetto stesso dell’analisi, sembra in ogni modo incrementata da alcune scelte tipografiche:<sup>5</sup>

- se “per ragioni tecniche le cartine, invece che direttamente nel testo corrente, hanno dovuto essere sistemate in blocco in fondo al volume” (6, nota 7), sarebbe stato quantomeno preferibile spostare in un’appendice sia la lista dei testi inseriti nel corpus, che l’elenco dei testi effettivamente analizzati, come anche il catalogo dei criteri, di cui, tra l’altro, solo una parte riceve una discussione più approfondita. Si tratta certamente di informazioni interessanti e indispensa-

---

<sup>3</sup> L’intero arco cronologico è suddiviso in cinque periodi: ‘I: anteriore al 1300; II: 1301-1350; III: 1351-1400; IV: 1401-1450; V: 1451-1525’ (17).

<sup>4</sup> La citazione è tratta da Goebel, Hans/Schiltz, Guillaume (2001): “Der ‘Atlas des formes et des constructions des chartes françaises du 13<sup>e</sup> siècle’ von Antonij Dees (1980) – dialektometrisch betrachtet”, in: Gärtner, Kurt *et al.* (a c. di): *Skripta, Schreiblandschaften und Standardisierungstendenzen*. Trier: Kliomedica, 169-221. Nel volume recensito essa appare, come tutte le altre citazioni prese da opere tedesche, tradotta in italiano (v. 419; ma con erronea indicazione di pagina (“170-171” invece di “170”).

<sup>5</sup> Considerando la mole e la complessità del libro, va sottolineato il bassissimo numero di errori tipografici, che solo eccezionalmente disturbano la lettura (tra l’altro: “ΣW<sub>4</sub>Corpus” (270, formula “a1”); a p. 313 mancano le sigle dei centri scrittorios (“2/[GE]”, ecc.); “i nessi lat. CE, CE [CI?]” (353); “6/UD TO” (414)). Segnaliamo infine che i rinvii bibliografici a “Videsott 2005” restano spesso ambigui tra “2005a” e “2005b” (v. 453).

bili, ma costituiscono grossi blocchi tipografici, la cui inserzione nel testo intralcia la lettura.<sup>6</sup>

- Accanto alle cartine, le tabelle sono certamente il modo più efficace per la presentazione dei tanti dettagli statistici emersi dalle analisi scrittologiche, ma la loro impostazione tipografica non risulta sempre felicissima: soprattutto nel caso l'estensione delle tabelle superi la lunghezza di una pagina e/o queste siano poste una di seguito all'altra, meglio sarebbe stato se le didascalie avessero preceduto le tabelle. In alcuni casi, la ripartizione su due pagine è solo dovuta allo spazio occupato dal testo che la precede e non alla lunghezza della tabella stessa; sarebbe stato generalmente preferibile evitare tali "interruzioni", che si rivelano particolarmente inconvenienti qualora i due segmenti della tabella finiscano sul fronte e sul retro di uno stesso foglio (v. ad esempio 309-310, 365-366, 403-404).
- Scomoda risulta anche la soluzione tipografica, adottata nelle analisi dettagliate relativamente ad alcuni dei criteri elencati, di apportare le prime e le ultime attestazioni del criterio in questione nel testo stesso, in forma più o meno estesa e senza messa in rilievo grafica dell'elemento che documenta il rispettivo fenomeno. Nel caso di "*De quel que sera trovaio delo meo en diner pagando la enpromesa de mia muier, la qual sé lib. M, voio qu'el sia sribuio per anema mia lib. cccc en cotal maniera 'denaro'*" (292), ad esempio, sarebbe bastato dire "*diner 'denaro'*" – soprattutto visto che, in seguito, le attestazioni vengono sistematicamente ripetute sotto forma di lista (v. ad esempio 293-296).

Pur non motivata nel suo complesso, la scelta dei criteri "analizzati a titolo esemplificativo" (290) sembra guidata soprattutto dal loro carattere costitutivo per la koiné padana (cfr. 311, 312, 332, 395) o per l'Italia settentrionale in generale (370), ma anche da altre ragioni.<sup>7</sup> Vorremmo però soffermarci su due altri criteri la cui discussione ci sembra di notevole pertinenza metodologica: scorrendo l'elenco dei 320 tratti ci si può stupire che la resa grafica del "nesso lat. tonico AL + cons." compaia come criterio sia a livello generale sia a livello "particolare", relativamente cioè alla resa grafica dei continuatori del lat. ALTER.<sup>8</sup> Come emerge però dal confronto delle relative analisi (v. 332-341<sup>9</sup> e 341-352), lo "sviluppo di AL + cons. [...] viene completato dalla parola ALTER in modo triplice" (352), visto che gli esiti di questa ultima presentano grafie ad essa particolari, hanno attestazioni ben più antiche e, inoltre, la resa grafica <a> è più ampiamente diffusa.<sup>10</sup>

---

<sup>6</sup> L'inconveniente è un po' mitigato dal segnalibro che accompagna il volume e che contiene i numeri, i nomi e le sigle dei rispettivi centri scrittori.

<sup>7</sup> Ad esempio la questione di una possibile conservazione prolungata di L nel nesso PL nell'Italia settentrionale, v. 382.

<sup>8</sup> V. 30, i criteri 23-26 (<Rivata> 'Rivalta', <fauz> 'falce', <Riolto> 'Rialto') e 30-31, i criteri 35-40 (<atre>, <oter>, <auto>, <altro>, <oltra>, <aytre>), rispettivamente.

<sup>9</sup> All'interno dell'analisi dedicata a questo criterio è stato preso in considerazione anche lo sviluppo del "nesso lat. tonico OL + cons." (30; v. 337-339).

<sup>10</sup> Un discorso analogo vale per il criterio 19, cioè la grafia <consa> 'cosa' <CAUSA (v. 329-331).

Sullo sfondo di quanto appena detto, appare opportuno ritornare in maniera più dettagliata sui criteri presi in considerazione dall'autore. Tralasciando le “ripetizioni”, che tali non sono che a prima vista,<sup>11</sup> sarebbe stato interessante sapere anche:

- qual è il rapporto tra i criteri 209 e 224, “r lat. non viene resa graficamente vs. tosc. <r>” e “Alle desinenze tosc. *-ari, -ori* di forme nominali al plurale corrispondono in pad. <ai>, <oi>”, nel caso delle occorrenze citate *segno* e *segnoi* ‘signor/signori’;
- come si giustifica la trattazione di pari passo delle varianti suffissali *-mentre, -menti* e *-mentri* ‘-mente’, cioè la loro ripartizione su tre criteri (231, 232, e 233);
- quali siano le eventuali ripercussioni sul calcolo soprattutto di  $Frel_{p/N/T/G}$ <sup>12</sup> dei criteri 159 e 161 o delle coppie 186/187 e 188/189, di cui è definitoria una restrizione diatopica, ad esempio: “Il nesso lat. CL in posizione intervocalica rom. è reso con <i> (*solo documenti liguri*) [...]” (40, criterio 159; corsivo L. F.) vs. “Il nesso lat. CL in posizione intervocalica rom. è reso con <j>, <i> (*in documenti non liguri*) [...]” (40, criterio 161; corsivo L. F.).

In una prospettiva più generale è necessario poi porre la questione del quadro di riferimento utilizzato per l'identificazione dei criteri da tenere in considerazione. La loro scelta è basata sulle “caratteristiche principali del sostrato dialettale generatore” (25), tratte da tutta una serie di studi (v. 25), mentre “[l]a presenza dei criteri è stata accertata tramite un confronto con l'italiano standard *moderno* (*contemporaneo*) e non in relazione al toscano medievale” (49); ne consegue ad esempio che occorrenze di <contrada> o di <lago> non sono state considerate, benché <contrata> e <laco> siano documentati in testi toscani medievali, e che è stato tenuto in considerazione, invece, il criterio ‘1ª pers. sg. ind. imp. in -a’, benché questa desinenza sia ampiamente documentata nel toscano medievale (v. 49). L'autore motiva “[q]uesta decisione a prima vista [...] anacronistica” con “un'ulteriore fase di ampliamento prevista per il nostro progetto, nella quale verrà inclusa anche l'analisi quantitativa dell'evoluzione del toscano verso la lingua standard odierna” (49). Tale ragionamento appare difficilmente ammissibile, non solo per motivi di principio – la questione dell'adeguatezza dei termini di paragone stessi non può essere risolta con il rinvio ad altre analisi –, ma anche per il fatto che non esiste un quadro di riferimento “toscano” omogeneo, né a livello cronologico né a livello diatopico (magari neanche a quello di tipi di testo). Sembra

<sup>11</sup> Nei criteri 294, 296, 298, 300, e 302 ad esempio compare la stessa formula descrittiva “Il nesso lat. -ĀTU è reso con [...]”, ma come emerge dagli esempi citati, nei primi quattro casi è questione di desinenze participiali, mentre nell'ultimo appare l'esempio *prò* ‘prato’. Segnaliamo comunque che per il criterio 296, “Il nesso lat. -ĀTU è reso con <at>, <ath> vs. tosc. <ato>”, l'esempio appor-  
tato non quadra: “*Ses avisat chu la chiasso chu ten Bartholomio di Pustiarmulo* ‘siete avvisato [sic]” (47); un problema simile riguarda il criterio 3, descritto come “-ĀTU/-ĀTU [sic; -ĀTI?]/-ATAE lat. sono resi con <é>, <eda> e <ede> vs. tosc. <ato>, <ata>, <ate>”, ma esemplificato con “*Miser Michel vardé ço che vui digé* ‘guardate’”: trattandosi di una forma all'imperativo, la desinenza latina è -ATE.

<sup>12</sup> Cioè il calcolo della “frequenza relativa” sull'intero corpus analizzato.

perciò anche ridotto il valore discriminativo di alcuni dei criteri presi in considerazione, dal momento che sono documentati anche in testi medievali toscani/fiorentini<sup>13</sup> fenomeni come ad esempio:

- la prostesi vocalica con *i-* davanti a *s* + consonante (v. 34, criterio 90<sup>14</sup>),
- sostantivi in *-ade/-ude* < lat. *-ATE(M)/-UTE(M)* (v. 37, criteri 141, 142),
- l'articolo *lo* (v. 43, criterio 242),
- *-amo*, *-emo* e *-imo* come desinenze della 1<sup>a</sup> pers. pl. ind. pres. dei verbi della 1<sup>a</sup>, della 2<sup>a</sup>/3<sup>a</sup> e della 4<sup>a</sup> coniugazione (v. 44, criteri 250, 259, e 45, criterio 265),
- il lessema *di* 'giorno' (v. 48, criterio 318).

Concludendo, non si può non condividere l'opinione espressa da Marcello Barbato: un lavoro "pionieristico in campo italiano" che "maneggia una grande quantità di dati, si confronta con numerosissimi problemi storico-linguistici e solleva interessanti questioni teoriche".<sup>15</sup> Queste ultime, lucidamente discusse dallo stesso Barbato, riguardano soprattutto il rapporto tra approccio scrittologico da un lato e filologia, dialettologia e storia della lingua dall'altro. Se tale rapporto non è certamente privo di problemi, ciò non toglie nulla al fatto che il libro di Videsott offre l'opportunità – ad entrambe le parti – di riflettere sui rispettivi presupposti metodologici.

Ludwig Fesenmeier  
Universität Erlangen-Nürnberg

---

<sup>13</sup> Tralasciamo in questa sede la variazione diatopica all'interno della Toscana stessa.

<sup>14</sup> Dove si parla di "epitesi".

<sup>15</sup> V. Barbato, Marcello (2010): "Scriptologia e filologia italiana. Accordi e disaccordi", in: *Medioevo romanzo* 34-1, 163-172; la citazione si trova a p. 167.

**Raffaella Bombi e Fabiana Fusco (a cura di) ... SAND CARRIED BY A STREAM... Scritti in onore di Vincenzo Orioles. Udine: Ed. Forum. 2009. 144 pp. ISBN: 978-88-8420-529-2**

*nel caro ricordo di Roberto Gusmani,  
nocchiero felice di una giornata  
difficile da dimenticare*

Quando, poco prima dell'estate, mi fu rivolto l'invito a presentare un volume di saggi in onore di Vincenzo Orioles, in occasione del suo sessantesimo compleanno, non ho avuto alcuna esitazione ad accogliere la proposta, e con grande piacere, misto a una nota malinconica, di cui pure dirò. La soddisfazione è certo suggerita da motivi scientifici, perché è sempre un grande onore poter festeggiare uno studioso largamente noto nel panorama della linguistica in Italia e all'estero, distintosi per una notevole varietà di interessi, ma nasce anche da ragioni personali, in quanto condivido con l'amico Orioles un lungo sodalizio, nato nei tempi ormai lontani - più di due decenni - del concorso che ci vide accomunati in un esito felice (e in ripetute visite a Viale Trastevere, aggiungerei, per far fronte ai continui intoppi burocratici che l'iter del decreto incontra). Il rammarico, anch'esso di natura personale, deriva dal fatto che quando si è invitati a presentare un volume miscelaneo significa che sono irrimediabilmente trascorsi gli anni della ricerca spensierata - si fa per dire, ovviamente -, e che sono arrivati tempi di più gravose responsabilità. Dato che si tratta, però, di un evento pressoché fisiologico, temo proprio che ci si debba rassegnare di buon animo.

Ma sarà meglio lasciare i preamboli, e venire al merito del bel volume, compatto e leggibile con interesse e profitto anche da chi non sia stretto praticante di discipline linguistiche e filologiche. La curiosità del lettore viene sollecitata immediatamente - com'è buona norma in questi casi - sin dal titolo, "... Sand carried by a stream...": la sabbia trasportata dalla corrente è la manifestazione metaforica dell'interferenza nell'atto - il mio Maestro avrebbe sottilmente usato l'univoco termine latino *actus* - di *parole*, secondo l'immagine adoperata dal Weinreich nel suo saggio più noto, *Languages in Contact*<sup>1</sup>. Ce lo ricordano le gentili Colleghe curatrici del bel volume, Raffaella Bombi e Fabiana Fusco, nella Premessa che segue la breve Presentazione ad opera del Direttore del Consorzio universitario del Friuli, Ernesto Liesch. Nella Premessa viene anche illustrato il senso dell'iniziativa, meritoria e al tempo stesso di grande impegno, in un momento nel quale il continuo mutare degli ordinamenti didattici e il fiorire di impegni accademici lasciano agli studiosi poco tempo per pensare e conseguentemente - almeno si spera - per scrivere. Si tratta, in sostanza, di un'opera

---

<sup>1</sup> Volume di cui il Festeggiato ha appena pubblicato una nuova edizione italiana (U. Weinreich, *Lingue in contatto*, Nuova ediz. italiana a cura di V. Orioles, con una Introduzione di G. R. Cardona, Torino, UTET, 2008), corredata di un'ampia Premessa che tra l'altro fa il punto sulla situazione degli studi sul contatto linguistico.

frutto di una volontà collettiva, che ha coinvolto i partecipanti, tutti Colleghi di sede del festeggiato, e accomunati da interessi scientifici omogenei, in un contributo, come si vedrà, al contempo articolato e convergente. Il titolo segnala non solo uno dei più ricchi àmbiti di ricerca di Vincenzo Orioles, l'interlinguistica, ma permette di riunire sotto un unico denominatore i sette articoli raccolti, che si inscrivono all'interno del più generale àmbito sociolinguistico (ciò vale anche, in misura maggiore o minore, per i due contributi di interesse ugristico).

Si è detto dei molti campi di ricerca coltivati dal festeggiato, e questa ricchezza di prospettive emerge con evidenza nelle 15 pagine che presentano la Bibliografia dal 1972 ad oggi: visto che sarebbe insostenibilmente scontato - benché sacrosanto - profondersi in espressioni di encomio sulla produzione scientifica di un Collega ed amico, trasferisco gli elogi sui curatori della bibliografia, che risulta estremamente accurata e affidabile, completa fino a comprendere tipologie di lavori che spesso vengono tralasciate in questo genere di raccolte, e di grandissimo vantaggio per chiunque voglia far ricerca nei settori indagati da Vincenzo Orioles. Questa sezione aggiunge senza dubbio ulteriore pregio a un volume che si distingue per la pulizia della stampa e per l'attenzione della cura editoriale (evidente anche nella rarità dei refusi tipografici).

Si è detto che la sociolinguistica costituisce in qualche modo il filo rosso che percorre la miscellanea, ma all'interno di questo ampio settore degli studi linguistici, le angolazioni e i problemi selezionati dai vari autori sono diversi, pur se in larga misura complementari.

Il tema sociolinguistico nel suo complesso è centrale nel saggio di Paolo Driussi, *La sociolinguistica oggi in Ungheria* (pp. 73-79): qui si ripercorre la storia degli studi sociolinguistici ungheresi, a partire dal fondatore, Miklós Kontra, epigono magiaro della scuola laboviana (cui appartenne sin dal periodo di lavoro oltre oceano). Un aspetto peculiare delle ricerche pubblicate in Ungheria in tale campo è rappresentato dal grande interesse per le dinamiche che regolano il rapporto tra la lingua magiara standard e le lingue delle minoranze ungheresi al di fuori dei confini politici attuali, in relazione alla forte posizione di prestigio dello standard rispetto alle varietà locali, spesso stigmatizzate (e su questo tema si innestano delicate questioni di politica linguistica, a partire dall'insegnamento, come ben evidenzia l'autore).

All'interno di una tematica fortemente improntata all'articolazione della lingua nella società si colloca anche il contributo di Fabiana Fusco (*Percorsi di parole 'al femminile': un sondaggio lessicografico*, pp. 81-115). Se certamente non è nuovo l'argomento da cui l'autrice trae spunto, quello del carattere più o meno sessista delle lingue - per obbligo istituzionale ricordo ad esempio, in Italia, il contributo di Anna Giacalone Ramat nel Convegno SIG del 1998<sup>2</sup> -, l'interesse del saggio, davvero consistente, è dato

---

<sup>2</sup> A. Giacalone Ramat, *Mutamento linguistico e fattori sociali: riflessioni tra presente e passato*, in *Linguistica storica e sociolinguistica*, Atti del Convegno della Società Italiana di Glottologia, Roma, Il Calamo, 2000, pp. 52-78.

dalla capacità di focalizzare l'attenzione sulla lessicografia in senso stretto, con una serie di considerazioni sostenute da un'analisi quantitativa riferita alla lingua italiana. Si scopre allora con qualche sorpresa come neppure un'opera certamente animata dalle migliori intenzioni riguardo alle "pari opportunità", il GRADIT demauriano, riesca a evitare le trappole - spesso nascoste - degli ideologemi sessisti che permeano di sé la storia della lessicografia italiana a partire per lo meno dal Tommaseo.

Una parte consistente degli articoli che compongono la Miscellanea concentra l'attenzione o comunque fa riferimento all'interferenza linguistica, quale elemento catalizzatore della evoluzione di lingue storiche del passato e del presente.

L'interlinguistica è la chiave che permette di interpretare una parte - più consistente di quel che potremmo immaginare - della lingua della politica italiana, come mostra nel suo corposo e argomentato saggio Raffaella Bombi (*Su alcune fonti alimentatrici esogene per la lingua speciale della politica italiana*, pp. 39-65). Molto spesso il modello straniero - per lo più anglo-americano - sfugge, vuoi perché mascherato - si pensi ai prestiti camuffati -, vuoi perché rielaborato nella forma più sottile dell'interferenza, quella del calco, spesso sintematico. Segnalo l'interesse particolare che riveste la discussione sulla distinzione tra calco semantico e prestito camuffato, tutt'altro che ovvia o banale. L'esemplificazione, particolarmente appropriata e di facile lettura, accresce il valore dell'attento lavoro di classificazione condotto nell'articolo.

L'interferenza è un tratto che - in positivo o in negativo - appare anche nella interpretazione storica di fatti linguistici - e culturali - fornita da altri tre degli autori che hanno contribuito al volume. Monica Ballerini offre una nuova analisi del toponimo *Castelbaldo*, che indica, in documenti che risalgono già alla fine del XIII secolo, una località in riva all'Adige, oggi in provincia di Padova. *Castelbaldo*, non c'è dubbio, è un composto il cui secondo elemento è germanico, ma dobbiamo dar credito alla tradizione che lo riporta al nome di un podestà patavino, appartenente alla famiglia dei Frescobaldi? L'autrice propone invece di vedere in *-baldo* l'equivalente di "forte", e allora *Castelbaldo* sarebbe più o meno corrispondente a *Castelforte*: una tale ipotesi potrà ovviamente trovare una conferma definitiva attraverso una scansione sistematica dei toponimi alto-italiani che presentano *-baldo* quale secondo elemento.

In un articolo riccamente documentato sul versante storico-culturale (*L'invasione hilaliana e la Storia della Sicilia: una piccola osservazione*, pp. 67-71), Guido Cifoletti pone in relazione la rapida - in certo senso sorprendentemente facile - conquista normanna della Sicilia con la complessa situazione dell'elemento arabo nell'isola e nella madre-patria tunisina. In particolare, l'interferenza manifestatasi in Tunisia proprio intorno alla metà dell'XI secolo tra l'elemento arabo preesistente, urbano e portatore di una *facies* linguistica evoluta (i cosiddetti dialetti pre-hilaliani), e l'elemento beduino proveniente dall'alto Egitto, di parlata hilaliana, dovette comportare tensioni anche socio-politiche, e dunque impedire qualunque tipo di sostegno attivo ai confratelli insulari (per conto loro tutt'altro che compatti di fronte agli invasori).

László Honti, in un denso studio (*Attributivkongruenz im Uralischen*, pp. 127-143), muove anch'egli da un fenomeno di interferenza, ma per smontare puntigliosamente,



uno ad uno, gli argomenti che lo sostengono. Si tratta della concordanza dell'attributo in numero e caso con il sostantivo, rintracciabile nel finnico e parzialmente in ungherese e in varie lingue uraliche, e che, secondo una interpretazione già risalente al Gabelentz, sarebbe da riportare a un influsso del germanico. L'indagine, condotta ad ampio raggio sull'intero gruppo delle lingue uraliche - nel quale viene compreso l'ugro-finnico -, si segnala per la documentazione e per il rigore; la proposta di spiegazione fatta propria dall'autore muove dall'osservazione relativa alla prevalente presenza della concordanza nelle costruzioni con il dimostrativo, che va interpretato come elemento focale della frase, mentre il sostantivo è di fatto una vera apposizione. Tale costruzione, in origine dunque apposizionale, viene poi intesa, in talune lingue, come attributiva, rappresentando infine il nucleo di espansione del tipo morfosintattico in questione.

Il magistrale saggio di Roberto Gusmani, intitolato *Lingua, cultura e caratteri genetici in un'ottica ricostruttiva*, pp. 117-126, occupa un posto a parte, ma non perché in qualche modo esuli dal tema del volume, ché anzi si tratta di contributo centrale e in certo senso basilare rispetto al rapporto tra linguistica storico-ricostruttiva e tratti socio-culturali - cui si aggiungono, in una discussione di grande impatto, i tratti genetici. Non mi soffermo sulla sezione dedicata alla ricostruzione, che costituisce la *summa* di un percorso di ricerca felicissimo condotto in vari decenni di studi da un Maestro insigne. Vorrei invece spendere qualche parola sul rapporto tra genetica e linguistica comparativo-ricostruttiva, argomento che aveva attirato l'attenzione anche del mio Maestro Walter Belardi, nei suoi ultimi anni di vita, e che ha visto alcuni importanti contributi nel convegno della Società Italiana di Glottologia a Verona nel 2007. Appare da sottoscrivere l'esclusione di qualsiasi parallelismo automatico tra *ethnos* e lingua, e finalmente trovo una esplicita negazione di legittimità alla menzione di un *Urvolk* indoeuropeo. E dunque il rifiuto di quella connessione tra lingua e patrimonio genetico, nel senso della dipendenza della prima dal secondo, incautamente sentenziata dalla scuola di Cavalli Sforza, offre un contributo fondamentale all'impostazione corretta del problema. Che appare suscettibile di soluzioni duttili, come già hanno suggerito altri genetisti, soprattutto Sokal e i suoi allievi, i quali hanno invertito il rapporto di dipendenza, facendo discendere le modifiche del patrimonio genetico dalle variazioni culturali e anche specificamente linguistiche, in una prospettiva che ciascuno di noi "glottosauri" si sentirebbe di poter accogliere senza problemi in base alla propria concreta esperienza di ricerca.

Come si può capire da questa rassegna inevitabilmente sintetica, il volume, pur dichiarando ambizioni modeste, ove si escluda il desiderio di festeggiare un peraltro ancor giovane Maestro, offre al lettore un quadro che non è affatto limitato, né per l'ampiezza dei temi - comunque coerenti con gli interessi coltivati dalla scuola linguistica udinese - né per lo spessore delle ricerche. La scelta di restringere gli inviti ai più stretti sodali di Vincenzo Orioles, come accennavo all'inizio, è stata in questo senso felicissima: un invito esteso all'intera linguistica italiana, o per lo meno a quella grande parte di essa che ha avuto rapporti di feconda collaborazione con la sede friulana, oltre a richiedere tempi lunghi, se non biblici, avrebbe quasi certamente dato luogo a

un ammasso entropico, ben difficile da ricondurre a una logica interna riconoscibile. Esattamente l'opposto di quel che possiamo constatare in questo caso: la "... Sand carried by a stream..." del titolo ha trovato la sua decantazione, e il risultato offre un saggio significativo e autorevole di quel settore di studi che - accanto a vari altri temi importanti, specie di linguistica storica - ha costituito un punto di forza, una sorta di marchio di fabbrica - se mi si permette l'espressione - della cerchia di linguisti raccolta intorno a Roberto Gusmani e Vincenzo Orioles.

Nell'avviarmi a concludere questa breve presentazione, mi sembra giusto rivolgere l'attenzione al festeggiato, per il quale il volume è stato concepito. Il fatto che in questa lieta circostanza chi vi parla si trovi a rappresentare la Società Italiana di Glottologia non è incidentale: a Vincenzo Orioles la Società Italiana di Glottologia deve molto, e l'amico e Collega ha ricoperto incarichi di responsabilità crescente, dapprima nel Direttivo, poi come Segretario, infine quale Presidente, in un momento non particolarmente facile per la linguistica in Italia - ma le sfide di allora oggi si rinnovano, in un quadro generale probabilmente ancor più fosco per l'Università nel suo complesso. Ecco, Enzo, in questo momento la Società Italiana di Glottologia per mio tramite non solo si unisce al festeggiamento, ma ti dice, con affettuosa semplicità: grazie!

Paolo Di Giovine  
Università di Roma "La Sapienza"

**Jan Radimský. *Les composés italiens actuels*. Paris : Cellule de Recherche en Linguistique. 270 pp. ISBN : 978-2-9526027-0-9**

Le travail de Jan Radimský, qui s'est donné pour objectif d'étudier les tendances récentes de la composition en italien, s'inscrit bien dans la thématique du présent numéro consacré aux frontières externes et internes de la morphologie. Bien que la composition constitue un des modes « centraux » de formation de mots, elle n'a, en ce qui concerne les langues romanes, jamais joui du même intérêt que la dérivation. Les raisons en sont multiples et ne sauraient être imputables à une « productivité » peu élevée des procédés de composition : comme le montre d'ailleurs l'analyse de J. Radimský, la disponibilité de la plupart des procédés de composition ne fait aucun doute et les composés néologiques recensés par les dictionnaires sont nombreux. Les raisons de ce relatif désintérêt résident plutôt dans l'hétérogénéité de ce mode de formation qu'il est difficile de décrire par une seule composante linguistique. Peut-on cerner les procédés de composition par des règles morphologiques ou bien relèvent-ils de la syntaxe ? Quelle y est la place de la composante lexicale et de la notion de « figement » ? Les mots composés doivent-ils être considérés comme des unités construites, selon la définition qu'en a donnée Danielle Corbin, ou bien s'agit-il simplement d'unités lexicales complexes ?

L'étude *Les composés italiens actuels* est une version revue et largement remaniée de la thèse de doctorat que l'auteur, actuellement enseignant-chercheur à l'Université de Bohême du Sud de České Budějovice (République tchèque), a soutenue à l'Université de Charles à Prague. Ajoutons encore que l'ouvrage a été publié par la Cellule de Recherche en Linguistique, une association scientifique, fondée par Amr Ibrahim (Université de Franche-Comté/Paris IV-Sorbonne) et Claire Martinot (Université Paris 5-René Descartes) et qui a pour objectif de promouvoir la recherche scientifique « fondamentale » en linguistique en organisant des colloques ou des journées d'études et en publiant des ouvrages spécialisés.

L'étude proposée par J. Radimský se base sur un corpus de 1390 néologismes, tirés du *Grande dizionario italiano dell'uso (Gradit)* de Tullio de Mauro (version électronique de 1999, Torino, Utet). Ce corpus, assez représentatif et fiable, regroupe les composés italiens dont la première attestation date de la période entre 1980 et 1999. Il convient de préciser qu'à la différence de la tradition des dictionnaires français, en règle générale plus conservateurs, les dictionnaires italiens recensent un nombre relativement grand de néologismes.

La première partie de l'ouvrage est consacrée à la définition des mots composés. L'auteur passe en revue la plupart des approches théoriques italiennes (Bologna, Tekavčić, Migliorini, Dardano, Scalise) et surtout françaises de la composition (Darmesteter, Guilbert, Corbin, G. Gross, Mathieu-Colas, Fradin). Cette « prépondérance » des linguistes français s'explique, d'une part, par la rareté des études théoriques détaillées pour l'italien et, d'autre part, par l'influence qu'ont exercée les études françaises de la question sur la morphologie italienne. De plus, l'approche de Jan

Radimský s'inscrit dans le cadre du structuralisme fonctionnel d'André Martinet qui n'a connu que peu d'écho dans la linguistique italienne. Au lieu de proposer une définition « en bonne et due forme » des composés et de la composition, ce qui s'avère impossible et soulèverait sans doute plus de problèmes que cela n'en résoudrait, l'auteur examine plusieurs types de critères : la soudure formelle des éléments (les degrés de l'identité phono-graphique du composé), les critères syntaxiques et morphosyntaxiques, les critères sémantiques et l'origine des éléments. Comme aucun des critères n'est ni nécessaire ni suffisant, J. Radimský envisage la catégorie des composés comme une catégorie prototypique et propose une « grille » qui permettra, d'un côté, de distinguer les composés des dérivés et des syntagmes libres et, d'autre côté, de déterminer la prototypicalité d'un composé ou d'un procédé de composition. Un composé prototypique devrait satisfaire aux critères suivants : soudure phono-graphique, distribution d'un lexème simple, non-actualisation des composants, autonomie des composants au moment de la fusion, blocage des transformations syntaxiques, absence de liaison syntaxique entre ses constituants, appartenance de sa structure à un modèle productif, signifié précis et unique, opacité sémantique. Une unité qui ne satisfera qu'à une partie des critères sera considérée comme un composé non prototypique. Si une unité ne répond à aucun des critères énumérés, elle est considérée comme un syntagme libre. Deux autres critères permettent de différencier un composé d'un dérivé : les constituants doivent appartenir aux parties du discours majeures et être dotés d'un sens référentiel.

Après un panorama rapide du traitement réservé aux composés italiens actuels par la linguistique contemporaine (S. Scalise, L. Serianni, G. Crocco Galèas et W. U. Dressler), quelques estimations sur l'importance quantitative des composés dans le lexique italien et une présentation détaillée du corpus, l'auteur passe à l'analyse de deux types de composés prototypiques (composés indigènes et confixés) et différents types de composition non prototypique, appelée anormale.

L'analyse des composés indigènes (it. *composto nativo*) actuels (chapitre V) permet de remettre en question un certain nombre d'observations antérieures. Si la composition est réputée être « essentiellement nominale » en italien, il est intéressant de voir une proportion non négligeable des adjectifs composés, plus de 15 %, auxquels on doit ajouter 17 % de composés connaissant à la fois un emploi adjectival et un emploi substantival. Selon S. Scalise un adjectif composé devrait avoir la structure interne (A+A), or, dans le corpus cette structure ne représente qu'un quart des adjectifs composés, la plupart comportant un élément verbal (V+N : *allungaciglia, spezzatimpani, portacasco, spruzzaprofumo* ; V+X : *paratutto, lavaindossa*). Les résultats montrent que les catégories considérées, selon S. Scalise, comme improductives ou indisponibles en italien actuel ont toutefois formé presque 15 % de tous les composés indigènes du corpus : il s'agit notamment des structures : A+N (*ambosessi, malagiustizia*), N+A (*postacelere, sassorosso*), P+N (*dopobagno, fuoripista*), V+X (*salvatutto, tagliacuci*). Il n'en est pas moins vrai que les modèles « centraux », à savoir V+N et N+N, restent les plus productifs aujourd'hui avec plus de 80 % des composés indigènes.

Le chapitre VI propose une analyse fouillée des confixés, qui manifestent en italien actuel, comme d'ailleurs dans la plupart des autres langues romanes, une très grande productivité, supérieure, d'après les statistiques établies par J. Radimský, à la productivité des composés indigènes. Le terme de « confixé », proposé par A. Martinet, remplace à juste titre le terme traditionnel de « composé savant » (it. *composto dotto, composto scientifico*). En effet, l'emploi de ces unités n'est plus limitée aux seuls registres « savants », de même que les composés indigènes ou populaires ne sont pas propres au seul registre commun : plus d'un quart des confixés du corpus dictionnaire ne portent aucune marque de spécialité, tandis que presque un tiers des composés indigènes, à l'en croire le GRADIT, relève de différents domaines de spécialité. De plus, une partie importante des confixes, comme nous le verrons quelques lignes plus loin, n'appartient pas au stock de formants d'origine gréco-latine.

La majeure partie de ce chapitre est consacrée aux traits typologiques des confixes. L'auteur insiste sur deux oppositions : préfixoïdes / suffixoïdes, confixes classiques / confixes modernes. La première opposition, dont les termes connaissent des acceptions fluctuantes en morphologie, réfère à la position que le confixe occupe dans le mot composé. La deuxième opposition repose essentiellement sur des critères étymologico-chronologiques et non pas fonctionnels ou structurels. Les confixes classiques sont des éléments de forme savante ou ayant subies des altérations phonétiques mineures, venant « directement » du stock gréco-latin. Les confixes modernes ou les néoconfixes présentent une catégorie relativement hétérogène englobant, d'un côté, les éléments issus de l'abréviation d'un lexème italien (*aviazione* > *avio-* > *aviosuperficie, aviocisterna*), ce qui génère parfois des homonymies (*auto-<sup>1</sup>* > *automobile / automobile* > *auto-<sup>2</sup>* > *autoambulanza*), et de l'autre, les éléments d'origine étrangère (*laser-* > *laserchirurgia, laserfoto*). Les néoconfixes, à la différence des confixes classiques, peuvent fonctionner comme des unités autonomes (*meteo, avio, porno, audio* etc.).

Les néoconfixes correspondent donc pour l'essentiel aux formants qui sont souvent considérés comme relevant de la morphologie « périphérique » ou même extragrammaticale : c'est le cas notamment de W. U. Dressler ou B. Fradin (composition cachée, affixation sécrétive). L'intégration de ces éléments à la catégorie des confixes s'explique par l'objectif de l'étude qui est d'établir les tendances de l'évolution actuelle de la composition, prise au sens large, en italien.

L'avant-dernier chapitre apporte l'analyse de formations extragrammaticales, dites « formations anormales », quantitativement peu importante (moins que 3 % du corpus), dont la plupart sont des mots-valises (it. *porola-macedonia*).

Le dernier chapitre, comme son titre l'indique, traite de « quelques problèmes résiduels ». Un premier concerne l'interférence entre la confixation et la composition indigène. Bien que le critère de l'ordre interne des éléments, qui correspond à « déterminé-déterminant » pour les composés et à « déterminant-déterminé » pour les confixés, reste pleinement opératoire, le fait que certains confixes connaissent un emploi autonome peut soulever des difficultés à distinguer nettement la confixation de la composition.

Il est intéressant que J. Radimský a décidé d'aborder le problème des composés à trois ou même à quatre formants dans ce chapitre « annexe » et non pas comme un type de formation anormale.

Après une analyse de quelques confixes (*mon-*, *bi-*, *multi-*), susceptibles d'opérer un changement de catégorie (N > Adj), l'auteur, en se basant sur l'analyse des néoconfixés de l'italien actuel, remet en question la pertinence du principe générativiste selon lequel la tête peut être identifiée par sa position dans le composé.

L'ouvrage de Jan Radimský, dont l'intérêt dépasse la morphologie de l'italien, est une tentative, intéressante et réussie, de proposer un traitement unifié de la composition en italien. Pour élargir la « portée » et vérifier les tendances observées à une plus grande échelle, il serait intéressant, d'un côté, de transposer, en l'adaptant, l'approche méthodologique et le cadre théorique à l'analyse de la composition dans d'autres langues romanes et, d'autre côté, d'effectuer une analyse analogue de composés néologiques non attestés (occasionalismes, hapax...).

Gregor Perko  
Université de Ljubljana

**Dany Amiot, Walter de Mulder, Estelle Moline et Dejan Stosic (éds.). *Ars Grammatica. Hommages à Nelly Flaux*. Bern: Peter Lang. 2011. 430 pp. ISBN: 978-3-0343-0691-1**

Le titre de ce recueil, publié en hommage à Nelly Flaux, est doublement évocateur : il rappelle, d'un côté, le domaine auquel Nelly Flaux a consacré plus de quarante ans de travail scientifique et universitaire et, de l'autre, le centre de recherche, Grammatica (EA 4521), qu'elle a créé à l'Université d'Arras en 1996. Grammatica est devenu l'un des centres les plus dynamiques de France dont les grands axes de recherche restent la syntaxe, la sémantique et l'analyse du discours. Le nombre de journées d'études ou de colloques organisés et d'ouvrages publiés ou dirigés par les membres du centre n'est que le signe le plus visible de ce dynamisme.

Les 24 articles qui composent l'ouvrage reflètent bien la diversité et la richesse thématique et théorique des recherches de N. Flaux. Il ne faut pas oublier que la syntaxe et la sémantique du nom (nom propre, différents types de noms communs) et du groupe nominal (détermination, quantification, pronoms), bien qu'occupant la position « centrale », ne constituaient pas ses seuls domaines de recherche, mais qu'elle s'intéressait également à la pragmatique, à la syntaxique, aux questions de l'aspect, des temps verbaux et à la pertinence de la notion de « manière » en linguistique. Il n'est donc pas surprenant que l'ouvrage soit organisé en trois parties : « Autour du nom et du groupe nominal », « De la syntaxe à la pragmatique » et « Autour du temps, de l'aspect et de la manière ».

L'article de Marie-Noël Gary-Prieur, qui ouvre ce recueil, propose une étude du « mot » *zéro*. Il s'inscrit dans le prolongement des recherches que l'auteur a amorcées déjà en 2004. *Zéro*, traditionnellement considéré comme un nom, est par certains aspects proche des numéraux, sans en faire partie, ce qui explique qu'il ait pu développer l'emploi en position de déterminant (*zéro faute*, *zéro stress*). Si son fonctionnement « originel » semble être bien celui d'un nom propre (emploi sans article, emploi en apposition, emploi en constructions « locatives »), il a développé des emplois propres au nom commun, essentiellement grâce aux mécanismes sémantiques, qui sont à l'origine de nombreux effets de discours.

Catherine Schnedecker s'intéresse à une catégorie « paradoxale » qu'elle appelle *noms propres « indéfinis »* (*Monsieur Tout-le-monde*, *Denise Trucmuche*, *Monsieur Durand* etc.). Tout en ayant, au niveau de la syntaxe et de l'orthographe, l'apparence de noms propres, ces noms se comportent, au niveau sémantique, comme des pronoms indéfinis (*quelqu'un*, *certain*).

L'article de Maria Ţenchea est la première des quatre contributions du volume rédigées par des collègues roumaines de N. Flaux. L'auteur se penche, dans une perspective contrastive, sur des aspects morphosyntaxiques et sémantiques de quelques adjectifs évaluatifs roumains (*bun*, *rău*, *frimos*, *urât*) en emploi nominal et leurs équivalents français.

La catégorie des adjectifs intéresse également Louis de Saussure. Il essaie d'établir des contraintes pesant sur la productivité dérivationnelle des adjectifs de couleurs. Si

les adjectifs désignant les couleurs focales admettent les dérivés d'approximation (*rouge* > *rougeâtre*) ou les dérivés verbaux (*rouge* > *rougir*), il n'en est pas de même pour les adjectifs désignant les couleurs non-focales (<sup>?</sup>*citronâtre*, <sup>?</sup>*mauvir*). L. de Saussure observe que les contraintes qui bloquent ces procédés dérivationnels ne sont pas de nature morphologique, mais plutôt d'ordre cognitif et pragmatique.

Dans l'article intitulé « Nommer Dieu, nommer les dieux à Rome », Françoise Toulze-Morisset montre que la différence entre la « nomination » *des* dieux dans la culture antique (Pline l'Ancien) et la nomination *de* Dieu dans la culture chrétienne ne peut pas être imputée à la différence entre la pluralité ou l'unicité du divin, mais à une conception radicalement différente de la nature même du divin.

Walter de Mulder et Anne Carlier engagent une discussion avec l'analyse de Michel Achard, d'inspiration cognitiviste et langackerienne, qui met en cause la distinction syntaxique entre les structures *il est/il arrive* et *c'est/ça arrive* et, par conséquent, la différence sémantique entre *il* et *ce/ça*. Les auteurs montrent qu'il convient de rétablir la distinction entre *il* impersonnel, référentiellement vide, et *ce* ou *ça* dotés d'un sens référentiel.

En s'appuyant sur la distinction entre le centre et la périphérie du système linguistique, concepts élaborés par l'École de Prague, et en transposant cette distinction des catégories elles-mêmes aux tendances de l'évolution diachronique, Bernard Combettes étudie la grammaticalisation et le problème de l'homogénéité des indéfinis en français.

Les deux derniers articles reviennent à la thématique de l'adjectif. Martine Coene et Liliane Tasmanowski traitent du double marquage de la définitude dans certaines langues balkaniques, notamment en roumain. L'explication qu'elles proposent pour l'emploi de *cel* en roumain, en distribution complémentaire avec l'accent tonique sur l'adjectif, procédé nettement plus fréquent en langue parlée, n'est pas d'ordre sémantique, mais s'appuie sur la notion de focus. Michele Prandi et Adriana Orlandi dressent un bilan des emplois figurés des adjectifs, insistant sur le rôle que jouent les propriétés grammaticales dans le déclenchement de l'interprétation figurée.

La deuxième partie, « De la syntaxe à la pragmatique », regroupe 10 articles. Les deux premiers proposent des études contrastives français-roumain. Eugenia Arjocaleremia analyse quatre cas de figure où l'équivalent roumain du pronom adverbial quantitatif *en* n'est pas exprimé (par exemple : fr. *on en trouve une quarantaine à Marseille* ; roum. : *La Marsilia Ø găsim vreo patruzeci*). Adina Tihu se penche sur les constructions de type *adj de N* dans lesquelles l'adjectif désigne une couleur (*brun de visage*, *vert de peur*, *vert de verdure*) et les constructions correspondantes en roumain.

Andrée Borillo examine les étapes et les conditions syntaxiques et sémantiques de l'évolution du nom *côté* qui peut en français actuel fonctionner comme une préposition (*côté jardin*, *côté confort*, *côté finances*). Elle avance l'hypothèse d'un processus de grammaticalisation passant par le figement de certaines structures prépositionnelles (*du côté de*). Cette hypothèse, comme le dit l'auteur, devrait être testée également sur d'autres « noms prépositionnels » (*genre*, *niveau*, *façon*, *question* etc.)



Marleen Van Peteghem se penche sur les conditions de l'interchangeabilité entre *tel (que)*, d'origine adjectivale, et *comme*, considéré comme un adverbe ou une conjonction, qu'elle analyse comme des proformes indéfinies.

L'examen du fonctionnement des prédicats verbaux causatifs de changement psychologique et la comparaison avec le fonctionnement des prédicats physiques ont permis à Danièle Van de Velde d'étendre la notion d'agentivité à tous les types de changements et de mettre en question la notion de « cause ».

Injoo Choi-Jonin s'interroge sur la pertinence du concept d'objet interne en français (par exemple, le groupe nominal *sa vie* dans *vivre sa vie*). Bien que certains linguistes jugent qu'il n'est pas justifié de distinguer l'objet interne de l'objet du verbe transitif pour les langues accusatives qui ne marquent pas, au niveau morphologique, la différence entre les deux types d'objets, l'auteur attire l'attention sur certaines propriétés syntaxiques que manifestent uniquement les objets internes et qui sont susceptibles d'assurer la « viabilité » de la distinction.

Les articles de Véronique Lagae et de Corina Cilianu-Lascu examinent deux « mots du discours », *à propos de* et *justement*. L'analyse des propriétés syntaxiques, sémantiques et discursives du marqueur *à propos de* en détachement frontal, effectuée par V. Lagae, montre qu'il faut distinguer deux emplois de ce marqueur : si *à propos de* est suivi d'un groupe nominal (par exemple, *à propos de son dernier roman*), il fonctionne comme un marqueur discursif, par contre, si *à propos de* ne s'accompagne pas d'un groupe nominal « plein », mais d'un nom ou d'un infinitif (par exemple, *à propos de sport*), il est plutôt un marqueur d'enchaînement discursif qui sert à reprendre un élément du contexte gauche et, ensuite, à réorienter le discours. S'appuyant sur un corpus bilingue, parallèle et comparable, d'une centaine d'exemples, C. Cilianu-Lascu étudie, dans une perspective contrastive, l'emploi pragmatique de *justement* et de ses équivalents roumains en position détachée.

L'article de Danielle Leeman et Céline Vaguer se penche sur la différence lexicale et sémantique entre *instant* et *moment* qu'on saisit intuitivement, mais qu'il est difficile d'explicitier. Elles défendent l'hypothèse, basée en partie sur l'étymologie, selon laquelle les deux mots s'opposent par la façon dont ils conceptualisent le temps.

Georges Kleiber et Marcel Vuillaume relancent le débat sur le paradoxe du répondeur (« the answering machine paradox ») qui a été lancée en 1991 par Alan Sidelle. Après avoir passé en revue les interprétations les plus importantes du problème soulevé par la petite phrase « *I am not here now* » (Predelli, Corazza, Fish et Gorvett, Voltolini, Stevens), ils concluent provisoirement (?) que l'essentiel du paradoxe réside dans la confusion entre le point de vue interne (de ceux qui participent à la communication) et le point de vue externe (du linguiste omniscient).

La dernière partie réunit les contributions qui s'intéressent à la problématique aspectuo-temporelle. Dans un article relativement court, mais très clair et riche, Carl Vettters s'est proposé de répondre à trois questions : Quelle est la fonction des temps verbaux ? Quelle est la nature de leur sens ? et Quels sont les pièges à éviter dans leur étude ? Il insiste sur la complexité des temps verbaux, tant au niveau de leur fonction-

nement (citons, entre autres, l'interaction avec le contexte, la présence de valeurs modales et aspectuelles qui se greffent ou s'ajoutent à la valeur temporelle de base) qu'en ce qui concerne leur sens (à la fois conceptuel et procédural). Il plaide pour une analyse « immanente » qui prend en compte tous les aspects de la langue, y compris sa dimension diachronique.

Les articles de Co Vet et de Laurent Gosselin traitent de la problématique de l'aspect. L'étude de Co Vet dresse des parallèles intéressants entre le sens des adverbies dits « aspectuels » *déjà* et (*pas*) *encore* et les aspects prospectif et résultatif et analyse les contraintes pesant sur la compatibilité de *déjà* avec un verbe au passé simple ou au futur simple. L. Gosselin propose un examen critique des deux approches de l'itération, approche dite nominale, qui a recours à la notion de quantification, et approche dite verbale, qui traite l'itération comme une catégorie aspectuelle, et des tentatives de rapprochement entre les deux auxquelles on assiste aujourd'hui. Dans sa conclusion, il met en garde contre un certain nombre de pièges qu'il convient d'éviter.

L'article d'Estelle Moline développe une piste suggérée par Nelly Flaux dans l'introduction au numéro thématique de *Langages* (n°175) consacré à l'expression de la manière. E. Moline s'intéresse aux « noms métalinguistiques » (*une façon de, une espèce de, un genre de*) et notamment à *une manière de*. Bien que ces constructions, en règle générale, n'imposent pas de contraintes sur le nom ou l'infinitif qui les complètent, elles manifestent clairement certaines préférences quant au choix du complément : par exemple, *une manière de* et *une façon de* se construisent essentiellement avec un infinitif, tandis que *une sorte de, une espèce de, un genre de* choisissent presque exclusivement les noms ; si *une manière de* et *une façon de* servent à établir uniquement des catégorisations de nature subjective, *une sorte de, une espèce de, un genre de* contribuent à établir en outre des catégorisations objectives et même scientifiques.

Avec le dernier article du recueil, on « renoue » avec la partie thématique du présent numéro de *Linguistica*. Dejan Stosic et Dany Amiot proposent d'étudier les verbes déverbaux évaluatifs formés par la suffixation (*bavardocher, gratouiller, trembloter*) au moyen de la notion de manière qui a été jusqu'alors réservée aux adverbies (en *-ment*), aux locutions adverbiales (*avec attention*) ou au sens codé lexicalement (*boiter*).

Ajoutons que l'ouvrage comporte une notice bibliographique et scientifique ainsi que la bibliographie complète de Nelly Flaux.

Gregor Perko  
Université de Ljubljana

## LINGUISTICA XLI-LI

### TABLE CHRONOLOGIQUE - KRONOLOŠKI PREGLED

#### XLI, 2001

1. Giuliano BONFANTE: *Pompetoxuaios*; 3-4
2. Janez OREŠNIK: *Naturalness expressed in sem-values - Jezikovna naravnost izražena v sem-vrednostih*; 5-20
3. Martina OŽBOT: *La coerenza testuale dal punto di vista traduttivo - Besedilna koherenca s stališča prevajanja*; 21-35
4. Gregor PERKO: *Les expressions idiomatiques: description théorique et traitement dictionnaire (l'exemple des dictionnaires monolingues français) - Idiomi: teoretični opis in slovarska obravnava (primer francoskih enojezičnih slovarjev)*; 37-75
5. Jerneja KAVČIČ: *On the use of the participle functioning as a finite verbal form in the Chronicle of John Malalas - Raba participa v funkciji osebne glagolske oblike v kroniki Janeza Malalasa*; 77-92
6. Leyre RUIZ DE ZAROBÉ: *Presupuestos lingüísticos para una enseñanza/aprendizaje del «discours rapporté» francés - Jezikoslovne postavke ob poučevanju/učenju odvisnega poročanega govora v francoščini*; 93-102
7. Pavao TEKAVČIČ: *Lingvistički pojmovi i izrazi u djelima Željke Čorak - Jezikoslovni pojmi in izrazi v delih Željke Čorak*; 103-110
8. Zorica VUČETIĆ: *La terminologia marinaresca studiata dal punto di vista della formazione delle parole - Pomorska terminologija s stališča besedne tvorbe*; 111-128
9. Paul VIDESOTT: *Gli italianismi nel ladino brissino-tirolese: alcuni aspetti quantitativi e cronologici in base all'ALD-I - Italianizmi v ladinščini južne Tirolske: gradivo iz jezikovnega atlasa dolomitske ladinščine*; 129-158
10. Varja CVETKO-OREŠNIK: *The o-grade of the Indo-European ablaut in the light of naturalness - Ojevaska stopnja praindoevropskega prevoja v luči jezikovne naravnosti*; 159-162
11. József Herman, *Vulgar Latin*, translated by Roger Wright, The Pennsylvania State University Press, University Park, Pennsylvania, XIV + 130 p. (Witold MAŃCZAK); 163-166
12. *Studi di linguistica greca II*, a cura di Emmanuele Banfi; Milano 1997, Francoangeli; *Materiali linguistici. Collana a cura dell'Università di Pavia, Dipartimento di Linguistica (Matjaž BABIČ)*; 167-176

13. Gramática descriptiva de la lengua española, dirigida por Ignacio Bosque y Violeta Demonte, publicada por la Real Academia española, colección Nebrija y Bello. Madrid: Espasa Calpe, 1999 (Jasmina MARKIČ); 176-178
14. Pavle Merku, Slovenska krajevna imena v Italiji, priložnik/Toponimi sloveni in Italia, manuale, Mladika, Trst 1999; pp. 126 (Mitja SKUBIC); 178-180

## XLII, 2002

15. Eric P. HAMP: Indic-nuristanī sákth-i/án-, čuṭti- and Albanian sup - Indijsko-nuristansko sákth-i/án-, čuṭti- in albansko sup; 3-4
16. Renato GENDRE: Nota gotica - Opomba k zgradbi sintagme v biblijski gotščini; 5-7
17. Sorin PALIGA: Herrscherschaft and herrschersuffix in Central-East European languages - Izrazi za gospodstvo in pripone zanje v srednjevhodnih evropskih jezikih; 9-18
18. Roxana IORDACHE: Le complément hypothétique en latin - Dopolnilo pogojnosti v latinščini; 19-29
19. Matjaž BABIČ: Les pronoms personnels ego et tu chez Plaute et Térence - Osebna zaimka ego in tu pri Plavtu in Terenciju; 31-42
20. Maria Rosaria CERASUOLO PERTUSI: Storie di parole ed etimi del dialetto triestino - Zgodovina in etimologija dveh besed iz triestinskega govora; 43-45
21. Tjaša MIKLIČ: Su alcuni usi tempo-aspettuali dei paradigmi verbali italiani trapassato prossimo e imperfetto – Nekatere temporalno-aspektualne rabe italijanskih glagolskih paradigem trapassato prossimo in imperfetto; 47-64
22. Zorica VUČETIĆ: Il linguaggio della giurisprudenza dal punto di vista della formazione delle parole: orientamenti e problemi lessicologici - Pravniški jezik v luči besedne tvorbe; 65-80
23. Vladimir POGAČNIK: L'imparfait et la modalité déontique - Francoski imperfekt in deontična modalnost; 81-88
24. Mojca SCHLAMBERGER BREZAR: Le rôle syntaxique et pragmatique des connecteurs dans le discours argumentatif français - Skladenjski in pragmatični vidiki povezovalcev v francoskih argumentacijskih besedilih; 89-110
25. Primož VITEZ: La norme démocratisée du français parlé - Demokratizirana norma francoskega govora; 111-120
26. Ana ZWITTER VITEZ: Les stratégies intonatives à l'échange oral en slovène et en français - Intonacijske strategije govorne izmenjave v slovenščini in francoščini; 121-130
27. Sonia VAUPOT: La détermination nominale: article et quantificateur - Samostalniška določnost: člen in količinski določevalnik; 131-142

28. Janez OREŠNIK: Naturalness in English: some (morpho)syntactic examples - Jezikovna naravnost v angleščini: nekaj (obliko)skladenjskih zgledov; 143-160
29. Gašper ILC, Milena MILOJEVIĆ SHEPPARD: Verb movement and interrogatives - Glagolski premik in vprašalni stavki; 161-176
30. Vida JESENŠEK: Okkasionelle Lexik in medialen Texten: pragmalinguistisch Betrachtet am Beispiel der Textsorte Pressekommentar - Priložnostno besedje v medijskih besedilih: pragmalingvistični vidiki rabe v besedilni vrsti časopisni komentar; 177-194
31. Ranko Matasović, Uvod u poredbenu lingvistiku. Zagreb: Matica Hrvatska, 2001. (Biblioteka Theoria, Novi niz); (Udžbenici Sveučilišta u Zagrebu = manualia Universitatis studiorum Zagrabiensis). 334 str. (Marina ZORMAN); 195-198
32. Jacqueline Picoche, Jean-Claude Rolland, Dictionnaire du français usuel. 15000 mots utiles en 442 articles. Bruxelles: De Boeck/Duculot. 2002 (deux éditions disponibles: édition papier et édition électronique sur cédérom PC/Mac) (Gregor PERKO); 198-202
33. Oana Sălișteanu Cristea, Introduzione alla dialettologia italiana, Tra lingua e dialetto, Editura DAIM, București 2002, 160 pagine (Pavao TEKAVČIĆ); 202-209
34. Claude Vincenot: Précis de grammaire logique. Honoré Champion, Paris, 1998 (Stojan BRAČIČ); 210-212
35. Nina Janich: Werbesprache. Ein Arbeitsbuch. Gunter Narr Verlag, Tübingen, 1999 (Stojan BRAČIČ); 212-214

### XLIII, 2003

36. Federico VICARIO: Note sull'ordine degli elementi in coppie di verbi antonimi - Pari pomensko nasprotujočih si glagolov: besedni red; 3-11
37. Matej HRIBERŠEK: Traditional theory about the origin of the Latin imperfect - Tradicionalna teorija o nastanku latinskega imperfekta; 13-24
38. Renato GENDRE: Sp. pescado – Kastiljsko pescado; 25-28
39. David BIZJAK: Verbo come elemento della frase in friulano ed in francese - Glagol kot prvina stavka v furlanščini in francoščini; 29-64
40. Mitja SKUBIC: Il pronome relativo che in friulano - Analitični oziralni zaimek v furlanščini; 65-72
41. Fiorenzo TOSO: Lo spazio linguistico còrso tra insularità e destino di frontiera - Jezikovna area korziščine med otoškostjo in usodnostjo mejnega govora; 73-92
42. Tjaša MIKLIČ: Il discorso indiretto libero nel romanzo di Giorgio Bassani Il giardino dei Finzi-Contini: funzioni testuali e caratteristiche linguistiche - Polpremi govor v romanu Il giardino di Finzi-Contini Giorgia Bassanija: besedilne funkcije in jezikovne značilnosti; 93-108

43. Vincenzo ORIOLES: Voci di origine russa in italiano - Besede ruskega izvora v italijanščini; 109-118
44. Janez OREŠNIK: Naturalness in English: (A) the genitive, (b) the pronouns - Jezikovna naravnost v angleščini - (A) rodilnik, (B) zaimki; 119-140
45. Tatjana MARVIN: Past participles in reduced relatives: a cross-linguistic prespective - Skladenjske lastnosti preteklega deležnika v deležniških polstavkih; 141-160
46. Goran Filipi, Istrorumunjski lingvistički atlas. Atlasul lingvistic istroromân. Atlante linguistico istroromeno, Knjižnica Atlas, Knjiga 2, Znanstvena udruga Mediteran=Societas studiorum Mediterraneum, Pula 2002, pag. 785 (Mitja SKUBIC); 161-163
47. Žarko Muljačić, Das dalmatische Studien zu einer untergegangenen Sprache, Quellen und Beiträge zur croatischen Kulturgeschichte, 10. Hg. Elisabeth von Erdmann-Pandžić, Böhlau Verlag, Köln Weimar Wien, 2000, pag. 434 (Mitja SKUBIC); 164-167
48. Snježana Kordić, Riječi na granici punoznačnosti. Zagreb: Hrvatska sveučilišna naklada, 2002, 227 str. (Marina ZORMAN); 168-172
49. Parallela, a.c. di/Hrsg. Roland Bauer e Hans Goebel, Testo - Variazione - Informatica / Text - Variation - Informatik, Atti del IX incontro italo-austriaco dei linguisti (Salisburgo, 1-4 novembre 2000) / Akten des IX. österreichisch-italienischen Linguistentreffens, (Salzburg, 1.-4. November 2000), Pro Lingua, Band 35; Gottfried Egert Verlag, Wilhelmsfeld 2002, 446 pp. (Pavao TEKAVČIĆ); 173-175
50. Dr. Roxana Iordache, Expimarea ideii de concesie în limba latină, Paideia, Colecția Academice, București 2002, 133 pagine (Pavao TEKAVČIĆ); 176-178

#### **XLIV, 2004**

51. Eric P. HAMP: Three brief studies: slovene and albanian - Tri kratke študije - slovenščina in albanščina; 3-6
52. MAGNÚS PÉTURSSON: Lautverschiebungen im Isländischen - Premiki glasov v islandščini; 7-16
53. Josip JERNEJ: Intorno all'analisi sintattica della frase semplice in italiano - K skladijski analizi prostega stavka v italijanščini; 17-26
54. Zorica VUČETIĆ: Il linguaggio della politica: il lessico della politica - Politični jezik. Politično besedje; 27-46
55. Martina OŽBOT: Tradurre per sbagliare, tradurre per imparare: la traduzione nello studio dell'italiano a livello universitario - Prevajanje kot sredstvo za usvajanje tujega jezika (primer študija italijanščine kot tujega jezika na univerzitetni stopnji); 47-58

56. Jerneja KAVČIČ: The greek infinitive in variable deliberative, principally dependent questions: an interpretation in terms of naturalness theory - Grški nedoločnik v vsebinskih deliberativnih, zlasti odvisnih, vprašalnih stavkih: razlaga s stališča teorije naravnosti; 59-74
57. Jacqueline OVEN: Le lexème slovène tudi et ses équivalents français - Slovenski členek tudi in njegovi ustrezniki v francoščini; 75-88
58. Meta LAH: Le type et l'authenticité du texte et leur influence sur la compréhension - Vpliv tipa in avtentičnosti besedila na njegovo razumevanje; 89-104
59. Gregor PERKO: Comment différencier les équivalents dans un dictionnaire bilingue d'encodage: (quelques propositions pour un dictionnaire slovène-français) - Kako razločevati prevodne ustreznice v dvojezičnem uvezovalnem slovarju (nekaj predlogov za slovensko-francoski slovar); 105-129
60. Alenka VRBINC, Marjeta VRBINC: How do Slovene learners of English use their dictionaries? - Kako rojeni govorniki slovenščine, ki se učijo angleščino, uporabljajo slovarje?; 131-160
61. Maurizio Puntin, Toponomastica storica del territorio di Monfalcone e del comune moderno di Sagrado, Centro Isontino di Ricerca e Documentazione Storica e Sociale »Leopoldo Gasparini«, Gradisca d'Isonzo - SKRD Jadro, Ronchi dei Legionari - SKŠRD Tržič, Monfalcone 2003, pag. 192 (Mitja SKUBIC); 161-166
62. Géographie linguistique et biologie du langage: autour de Jules Gilliéron. Édité par Peter Lauwers, Marie-Rose Simoni-Aurembou, Pierre Swiggers, Orbis supplementa t. 20. Leuven - Paris - Dudley, MA, Peeters Leuven 2002; 212 pp. (Mitja SKUBIC); 167-171
63. Grammaire et enseignement du français, 1500-1700. Édité par Jan de Clercq, Nico Lioce, Pierre Swiggers; Monographie publiée par le Centre international de dialectologie générale (Louvain) ; Peeters, Leuven-Paris-Sterling, Virginia, 2000, XXXIV+671 pp. (Orbis/Supplementa, t. 16) (Vladimir POGAČNIK); 172-173
64. Syntax in Antiquity, edd. P. Swiggers - A. Wouters; Louvain 2003, Peeters (Orbis-Supplementa t. 23) (Matjaž BABIČ); 174-178
65. Roxana Iordache: Studii de lingvistică latină și lingvistică romanică, Editura BREN, București 2002, 144 pp. (Pavao TEKAVČIČ); 179-180
66. Dr. Roxana Iordache, Aspecte ale coordonării în limba latină, Editura DAIM, București 2003, 161 pp. (Pavao TEKAVČIČ); 181
67. Walter Belardi, Breve storia della lingua e della letteratura ladina, 2. edizione aggiornata; con un'appendice di Marco Forni; Istitut Ladin »Micurà de Rù«, San Martin de Tor, 2003; 138 pp.+indice (5 pagine non numerate) (Pavao TEKAVČIČ); 182-183
68. Giuseppe Patota, Grammatica di riferimento della lingua italiana per stranieri, Società Dante Alighieri; Firenze, Le Monnier, 2003; 424 pp. (Pavao TEKAVČIČ); 184-187

69. Quaderni di filologia e lingue romanze, Ricerche svolte nell'Università di Macerata, Terza serie, vol. 17; Macerata 2002, 414 pp. (Pavao TEKAVČIĆ); 188-189
70. Il Registro della confraternità dei pellicciai di Udine. A cura di Federico Vicario. Biblioteca di Lingua e Letteratura friulana, 4. - Forum, Editrice Universitaria Udinese, Udine 2003; pp. 204 (Mitja SKUBIC); 190-192

#### XLV, 2005

71. Marija GOLDEN: Janez Orešnik on his seventieth birthday; 3-5
72. Tabula gratulatoria; 6-8
73. Ana JUVANČIČ-MEHLE: Bibliographie sélectionnée des oeuvres du prof. Janez Orešnik - Izbrana bibliografija del prof. Janeza Orešnika; 9-28
74. Wolfgang U. DRESSLER: Towards a natural morphology of compounding - Prispevek k naravnemu oblikoslovju zloženk; 29-40
75. Pierre SWIGGERS: Formes, fonctions et catégories linguistiques: Les principes de grammaire générale de Louis Hjelmslev - Oblike, funkcije in kategorije v jeziku: Principes de grammaire générale Louisa Hjelmsleva; 41-52
76. Kenneth SHIELDS: A proposal about the origin of the Indo-european locative plural - Nova domena o izvoru indoevropskega mestnika množine; 53-58
77. Toshiko YAMAGUCHI, Magnús PÉTURSSON: Das Isländische Hafa-perfekt in Zeitungsartikeln - Islandski perfekt s pomožnikom hafa v časopisnih člankih; 59-84
78. John WEINSTOCK: The bear went down the mountain: Sámi language usage then and now - Medved se je spustil v dolino - raba samijskega jezika nekdanj in zdaj; 85-96
79. Tatjana MARVIN: English syllabification and schwa-insertion: from The sound pattern of English to the notion of phase - Interakcija med zlogovanjem in vrivanjem polglasnika v angleščini z vidika dela »The sound pattern of English« ter teorije faznega pristopa; 97-106
80. Gašper ILC, Milena MILOJEVIĆ SHEPPARD: Negative pronouns in Slovene: a minimalist account - Nikalni zaimki v slovenščini: minimalistična razčlemba; 107-120
81. Witold MAŃCZAK: Origine du vieux slave - Izvor stare cerkvene slovanščine; 121-128
82. Heinz Dieter POHL: Die Slavia submersa in Österreich: ein Überblick und Versuch einer Neubewertung - Slavia submersa v Avstriji: pregled in poskus pre vrednotenja; 129-150
83. Jože TOPORIŠIČ: Naglasna znamenja v Bohoričevih Zimskih uricah 1584 – Die Betonungszeichensetzung in Bohorič's Arcticae horulae 1584; 151-158
84. Erika KRŽIŠNIK: A type of minimal phraseological units characteristic of Slovene – Značilen tip slovenskih minimalnih frazemov; 159-172
85. Florence GACOIN-MARKS: Vladimir Levstik et la langue française avant 1941 - Vladimir Levstik in francoski jezik pred letom 1941; 173-182



86. Pavao TEKAVČIĆ: O tzv. pučkoj etimologiji (tj. motivaciji) - Motivacija v jeziku (ljudska etimologija); 183-188
87. Žarko MULJAČIĆ: Nuovi dati sulle colonie croate nell'Italia meridionale con particolare riguardo a quelle estintesi da secoli in Campania - Novi podatki o hrvaških naselbinah v južni Italiji; 189-202
88. Jerneja KAVČIČ: Pleonastic use of verbs of speaking in Greek: an interpretation in terms of naturalness theory - Pleonastična raba glagolov govorenja v grščini: razlaga s stališča teorije naravnosti; 203-212
89. Roxana IORDACHE: La subordonnée conditionnelle dans le bas latin, le registre populaire - Pogojni odvisniki ljudske rabe v pozni latinščini; 213-224
90. Matjaž BABIČ: Word order variation in Plautus - Besednoredne inačice pri Plavtu; 225-238
91. Tjaša MIKLIČ: » Flash-forward » in italiano: aspetti concettuali e moduli espressivi - Flash-forward v italijanščini: pojmovni vidiki in izrazni moduli; 239-258
92. Fiorenzo TOSO: L'avverbio e pronomi ghi in dialetti corsi e peri-corsi - Prislov in zaimki ghi v korziških osrednjih in obrobnih narečjih; 259-276
93. David BIZJAK: Quale sintagma usare per indicare la necessità dell'azione in friulano - Sintagme za izraz potrebe po dejanju v furlanščini; 277-282
94. Mitja SKUBIC: Il dativo etico nel friulano sonziaco - Etični dativ v obsoški furlanščini; 283-288

#### XLVI, 2006

95. Hans GOEBL: Sur le changement macrolinguistique survenu entre 1300 et 1900 dans le domaine d'oïl: une étude diachronique d'inspiration dialectométrique - Jezikovne inovacije med 14. in 16. stoletjem na ozemlju francoščine. Diahronična študija z vidika dialektometrije; 3-43
96. Marek GAWĘŁKO: Sur un aspect des études historiques - Razvojni vidik v indoevropskih jezikih; 45-54
97. Zuzanna TOPOLINJSKA: Toward a definition of the generic noun phrase - K definiciji t.i. generične imenske sintagme; 55-59
98. Lara BURAZER: Referring and the context of situation in view of the human linguistics approach - Pojmovanje reference (nanašanja) in situacijskega konteksta s stališča jezikoslovja človeka; 61-70
99. Martina OŽBOT: On the possibilities and impossibilities of cross-cultural communication - O možnostih in nemožnostih medkulturne komunikacije; 71-80
100. Karmen TERŽAN-KOPECKY: Translate als kognitive Entitäten oder die Natürlichkeitstheorie im Kontext der sprachlichen und psychosozialen Paradigmata der mehrsprachlichen Kommunikation - Translati kot spoznavne

- entitete ali teorija naravnosti v kontekstu jezikovnih in psihosocialnih paradigem večjezikovnega sporazumevanja; 81-98
101. Teodor PETRIČ: Vorschläge zur Erweiterung der slowenischen Variante der Natürlichkeitstheorie - Predlogi za razširitev slovenske različice teorije naravnosti; 99-115
  102. Varja CVETKO-OREŠNIK: The aorist and the perfect of the Old Indian causatives in the light of natural morphosyntax - Aorist in perfekt staroindijskih kavzativov v luči naravne oblikoskladnje; 117-120
  103. Zorica VUČETIĆ: Il linguaggio artistico: il lessico - Jezik umetnosti. Besedje; 121-144
  104. Thomas L. MARKEY: Early celticity in Slovenia and at rhaetic Magrè (Schio) - Zgodnja keltščina v Sloveniji in retskem kraju Magrè (Schio); 145-172
  105. Irena OREL: Prepositional phrases with verba dicendi from Dalmatin's translation of the Bible (1584) in relation to foreign language translations - Predložne zveze ob glagolih govorjenja iz Dalmatinovega prevoda Biblije (1584) v razmerju s tujejezičnimi prevodi; 173-180
  106. Alja FERME: Final sonorant sequences in the Celje dialect - Zvočniška izglasja v celjskem govoru; 181-189
  107. Chikako SHIGEMORI BUČAR: Causative constructions in japanese and slovene - Vzročniške zgradbe v japonščini in slovenščini; 191-202
  108. Aleksandra DERGANČ: Trije zgledi neposrečenega prevoda zaradi »lažnih prijateljev« - Three instances of incorrect translations due to false friends; 203-208
  109. Eva SICHERL, Andreja ŽELE: Combinability with prepositions in Slovene and English: theoretical and contrastive views - Predložne zveze oz. predložnomorfemskost v slovenščini in angleščini: teoretični in kontrastivni vidik; 209-226
  110. Marjeta VRBINC, Alenka VRBINC: A research-based study of foreign students' use of grammatical codes in five leading British learners' dictionaries - Raziskava o uporabi slovničnih kod v petih vodilnih britanskih enojezičnih slovarjih za tujce na vzorcu tujih govorcev; 227-242
  111. Jasmina MARKIČ: Valores y usos de las perífrasis verbales de gerundio con los auxiliares ir, andar y venir - Vrednosti in rabe glagolskih perifraz s pomožniki ir, andar in venir v gerundiju; 243-250
  112. Vladimir POGAČNIK: Nouvelle contribution à l'étymologie du toponyme Ljubljana - Nov prispevek k etimologiji toponima Ljubljana; 251-253

## XLVII, 2007

113. Mitja SKUBIC: Pavao Tekavčić (1931-2007); 3-4
114. Pavao TEKAVČIĆ: O smrti latinskog jezika i nekim s tim povezanim temama: osvrt na knjigu Halte à la mort des langues Claudea Hagègea - Il latino, la »morte« o la continuità; 5-9

115. Gabriele Iannàcero [i. e. Iannàcaro], Il dialetto percepito, Sulla reazione di parlanti di fronte al cambio linguistico; *Lingua, cultura, territorio*, Collana diretta da Tullio Telmon 30; Edizioni dell'Orso, Alessandria 2002; 250 pagine (Pavao TEKAVČIĆ): 10-11
116. Académie Roumaine, Institut de Linguistique »Iorgu Iordan - Al. Rosetti« - Université de Bucarest, Faculté de Langues et Littératures Étrangères, Dictionnaire des emprunts latins dans les langues romanes, Jana Balacciu Matei ... ; sous la direction de Sanda Reinheimer Rîpeanu; Editură Academiei Române, București 2004; 456 pagine (Pavao TEKAVČIĆ); 12-13
117. Marcello Aprile, Le strutture del lessico etimologico italiano; pubblicazioni del Dipartimento Etimologico Italiano; pubblicazioni del Dipartimento di Filologia Linguistica e Letteratura dell'Università di Lecce, vol. 24; Mario Congedo editore, Galatina 2004; 324 pagine (Pavao TEKAVČIĆ); 14-16
118. Giuseppe Patota, Poiché fra causa, tempo e testo, *Biblioteca di cultura / 675*, Bulzoni editore, Roma 2005, 398 pp. (Pavao TEKAVČIĆ); 16-17
119. Michele Prandi, Gaston Gross, Cristiana De Santis, La finalità, strutture concettuali e forme d'espressione in italiano; *Biblioteca dell'Archivum Romanicum, Serie II: Linguistica*, 56; Leo S. Olschki editore 2005; 364 pagine (Pavao TEKAVČIĆ); 18-20
120. Pierre SWIGGERS: Two key concepts of language endangerment: language obsolescence and language death - Dva ključna pojma nevarnosti v jeziku: zastarelost in izginotje; 21-33
121. Janez OREŠNIK: Natural syntax: English interrogative main clauses - Naravna skladnja: angleški vprašalni glavni stavki; 35-48
122. Concepción CABRILLANA: Definiteness strategies and word order in existential-locatives and locatives in Late and Vulgar Latin – Določnostne strategije v besednem redu in v eksistencialno-lokativnih zvezah v pozni latinščini; 49-64
123. Tjaša MIKLIČ, Martina OŽBOT: Teaching the uses of Italian verb forms to Slovene speakers - Poučevanje rab italijanskih glagolskih paradigem pri govorcih slovenščine; 65-76
124. Eva NÚÑEZ: ¿Tienen ventajas cognitivas los bilingües sobre los monolingües? - Domnevne prednosti dvojezičnosti pred enojezičnostjo; 77-85
125. Meta LAH: Le processus de la planification lors de l'écriture en langue étrangère: la langue et la structure - Proces načrtovanja pri pisanju v tujem jeziku: jezik in struktura; 87-100
126. Nataša ŽUGELJ: Les erreurs des étudiants slovènes dans les productions écrites - Napake slovenskih študentov francoščine v pisnih sestavkih; 101-108
127. Mojca SMOLEJ: La classification des particules en français et en slovène - Obravnava členkov v francoskem in slovenskem jezikoslovju; 109-120

128. MAGNÚS PÉTURSSON: Grundlagen für phonetische Übungen mit isländisch-sprechenden - Podlage za fonetične vaje pod vodstvom rojenih govorcev islandščine; 121-136
129. Ada GRUNTAR JERMOL: Fachsprachliche Kollokationen: Schwerpunkt Deutsch-Slowenisch - Kolokacije s težiščem na nemško-slovenskem strokovnem jeziku; 137-144
130. Alenka VRBINC, Marjeta VRBINC: EFL students' use of grammatical information in five leading British learners' dictionaries - Kako govorniki angleščine kot tujega jezika uporabljajo slovnične informacije v petih vodilnih britanskih slovarjih za tujce; 145-158
131. Un importante contributo alla lessicografia italo-slovena: Sergij Šlenc, Veliki slovensko-italijanski slovar - Il grande dizionario sloveno-italiano, Državna založba Slovenije, Ljubljana, 2006; pp. XIV + 1539 (Mitja SKUBIC); 159-165
132. Rada Cossutta - Franco Crevatin, Slovenski dialektološki leksikalni atlas slovenske Istre (SDLA-SI) II, Založba Annales 2006 (Mitja SKUBIC); 166-167
133. Mons. Ivan Trinko (1863-1954). Spodbujevalec spoznavanja in dialoga med kulturami/Promotore della conoscenza e del dialogo tra le culture. Kulturno društvo/Circolo culturale I. Trinko, Čedad/Cividale 2006 (Mitja SKUBIC); 168-169
134. Il dialetto tedesco dell'isola alloglotta di Sappada, tesi di laurea [di] Maria Bruniera (Darko ČUDEN); 170-171

#### XLVIII, 2008

135. Martina OŽBOT: Il professor Mitja Skubic ottuagenario; 7-8
136. Tabula gratulatoria; 9
137. Marjeta PRELESNIK DROZG: Bibliografija prof. Mitja Skubica (1952-2008) - Bibliographie du professeur Mitja Skubic (1952-2008); 11-31
138. Janez OREŠNIK: Standard French liaison and natural syntax - Vežanje (liaison) v knjižni francoščini in naravna skladnja; 33-48
139. Ana ZWITTER VITEZ: Principes narratifs dans l'échange oral spontané et leurs spécificités en slovène et en français - Narativna pravila spontane govorne izmenjave in njihove specifičnosti v slovenščini in francoščini; 49-58
140. Meta LAH: Le fait divers, une narration défailante? - Članek iz (črne) kronike, pripoved z napako?; 59-72
141. Gregor PERKO: Les je gigognes du roman célinien - Jaz znotraj jaza v romanih L.-F. Célinea; 73-82
142. Primož VITEZ: L'invention du texte didascalique - Didaskalija kot besedilo; 83-94
143. Sonia VAUPOT: Narration et langage dans l'oeuvre de Raymond Queneau, Zazie dans le métro - Naracija in jezik v Queneaujevem delu Cica v metroju; 95-101

144. Tjaša MIKLIČ: Alcuni aspetti di tipo pragmatico e di quello retorico-narratologico nell'uso del sistema verbale italiano: con possibili implicazioni per l'insegnamento L2 - Nekateri vidiki pragmatične oz. retorično-naratološke narave v rabi italijanskega glagolskega sistema (z možno navezavo na poučevanje drugega jezika); 103-120
145. Darja MERTELJ: Frasi all'imperativo in italiano: aspetti glottodidattici di forma, intonazione, cortesia ed interculturalità - Velelne povedi v italijanščini (oblikovni, intonacijski, vljudnostni in medkulturni vidiki v luči poučevanja jezika); 121-135
146. Jana KENDA: Ruoli pragmatici di »si«: modello di analisi dei segnali discorsivi nell'italiano parlato - Pragmatične vloge »si«: model analize diskurznihih označevalcev v govorjenem italijanskem jeziku; 137-149
147. Agata ŠEGA: Alcune caratteristiche dello sloveno occidentale tra lingua e parola - Nekaj značilnosti zahodne slovenščine med jezikom in govorom; 151-158
148. Martina OŽBOT: Alcuni cenni sugli italianismi in sloveno - Nekaj opazanj o italianizmih v slovenščini; 159-166
149. Jasmina MARKIČ: Análisis textual de la narrativa oral de los Llanos Colombianos: ejemplo de un informante de Casanare, Colombia - Besedilna analiza ustne pripovedi kolumbijskih Llanerosov: primer informanta iz departmaja Casanare; 167-176
150. Barbara PIHLER: El discurso poético de Juan Ramón Jiménez: eternidades en espacio (enfoque pragmático) - Pesniški diskurz španskega pesnika Juana Ramóna Jiménez: večnost v prostoru (pragmatični vidik); 177-189
151. Branka KALENIĆ RAMŠAK: El Quijote de dos caras: Cervantes versus Avellaneda - Don Kihot z dvema obrazoma: Cervantes proti Avellanedi; 191-204
152. Maja ŠABEC: La Celestina, »novela dialogada« y »novela dialógica« - Celestina, »dialogizirani« in »dialoški roman«; 205-214
153. Smilja KOMAR: Intonation, the strongest link in the Weakest link quiz - Intonacija, najmočnejši člen v kvizu Najšibkejši člen; 215-226
154. Silvana OREL KOS: A contrastive-stylistic study into the tense distribution in English and Slovene fictional texts - Razporeditev glagolskih časov v angleških in slovenskih pripovednih besedilih s kontrastivno-stilističnega stališča; 227-236
155. Stojan BRAČIČ: Semantische Merkmale als Mittel der Textkonstitution - Semantični znaki kot sredstvo besedilne konstitucije; 237-244
156. Janja POLAJNAR: Zur Explizierung von Adressierungsstrategien in deutschen Kinderwerbepots - O ekspliciranju strategij naslavljanja v nemških oglasih za otroške izdelke; 245-258
157. Andreja ŽELE, Eva SICHERL: Präfixal-präpositionale Verhältnisse bei Slowenischen präfigierten Verben - kontrastiert mit dem Deutschen - Predponsko-predložna razmerja v slovenščini - kontrastirano z nemščino; 259-276

158. Andrej BEKEŠ: Japanese suppositional adverbs: probability and structure in speaker-hearer interaction - Ugibalni prislovi v japonščini: verjetnost in struktura v interakciji med govorcem in sogovorcem; 277-292
159. Metka FURLAN: The origin of Anatolian relations of the type keššar : kiššeran and Balto-Slavic relations of the type akmuo/kamy: akmeni/kamen[jer] - O nastanku anatolskih razmerij tipa keššar : kiššeran in balto-slovanskih tipa akmuo/kamy : akmeni/kamen[jer]; 293-307

#### XLIX, 2009

160. Martina OŽBOT: Nota redazionale; 7
161. Gaetano BERRUTO: Sul posto della variazione nella teoria linguistica - O mestu, ki pripada jezikovni raznolikosti v teoriji jezika, 9-25
162. Augusto CARLI: Per un aggiornamento del concetto di vitalità linguistica - K posodobitvi koncepta jezikovne vitalnosti; 27-39
163. Paola DESIDERI: Origini e sviluppi delle analisi e delle teorie sul linguaggio politico (1920-1960) - Izvor in razvoj analiz in teorij političnega jezika (1920-1960); 41-53
164. MAGNÚS PÉTURSSON: Warum sollen wir Sprachen lernen und erforschen? - Zakaj se moramo učiti jezikov in jih raziskovati?; 55-64
165. Janez OREŠNIK: Transitivity in natural syntax: ergative languages - Prehodnost v naravni skladnji: ergativni jeziki; 65-93
166. Michele A. CORTELAZZO: Evoluzione della lingua e staticità della norma nell'italiano contemporaneo: gli ausiliari nei costrutti con verbi modali - Razvoj jezika in statičnost norme v sodobni italijanščini: pomožniki v konstrukcijah z modalnimi glagoli; 95-106
167. Gunver SKYTTE: La connessione »implicita« - »Implicitna« zveza; 107-120
168. Rosanna SORNICOLA: I tipi italiani conesso, lunghesso, sovrasso, sottesso e la grammaticalizzazione di ipse - Italijanske zveze conesso, lunghesso, sovrasso, sottesso in gramatikalizacija zaimka ipse; 121-130
169. Sandra AUGENDRE: Soggetto e verbo con dormire, telefonare e arrivare in contesti diversi - Osebek in glagol z dormire, telefonare in arrivare v različnih kontekstih; 131-159
170. Mirjam PREMRL: Commento delle scelte delle forme verbali nei cloze degli studenti di Italianistica e di Traduzione (identificazione dei punti critici e confronto tra i gruppi) - Komentar k izbiri glagolskih oblik v clozih študentov Italijanistike in Prevajanja (opredelitev kritičnih točk in primerjava posameznih skupin); 161-203
171. Fabiana FUSCO: Stereotipo e genere: il punto di vista della lessicografia - Stereotip in spol: leksikografski vidik; 205-225

172. Vincenzo ORIOLES: Tra sicilianità e sicilitudine - Med sicilijanstvom (sicilianità) in sicilijanskostjo (sicilitudine); 227-234
173. Fiorenzo TOSO: Parole ritrovate - Znova najdene besede; 235-245
174. Alberto ZAMBONI: Nuove osservazioni su una glossa botanica mediolatina: citamus 'colchico' - Nova opažanja o srednjelatski botanični besedi: citamus 'podleseki'; 247-256
175. Roberto GUSMANI: Il toponimo friulano Nimis - Furlanski toponim Nimis; 257-260
176. Federico VICARIO: Il catalogo di donne di Venzone (sec. XIV) – Il catalogo di donne di Venzone (14. stol.); 261-276
177. Roberto DAPIT: Relazioni semantiche tra lo sloveno standard e i dialetti con riferimento alle lingue di interazione - Pomenska razmerja med standardno slovenščino, njenimi narečji in interakcijskimi jeziki; 277-293
178. Liliana SPINOZZI MONAI: Ipotesi di un calco paradigmatico slavo-romanzo (l'imperativo-congiuntivo, uno studio fondato sul Glossario del dialetto del Torre di Jan Baudouin de Courtenay) - Hipoteza o paradigmatškem slovansko-romanskem kalku (imperativ-konjunktiv: raziskava, temelječa na Slovarju terskega narečja Jana Baudouina de Courtenayja); 295-308

## L, 2010

179. Roland BAUER: Die Position des Rätoromanischen und seine Beziehungen zum Deutschen, Französischen und Italienischen - Položaj retoromanščine in njeno razmerje do nemščine, francoščine in italijanščine; 7-26
180. Hans GOEBL: Del posto dialettometrico che spetta ai punti galloitalici Aidone, Sperlinga e San Fratello nel sistema della rete dell' AIS - O dialektometričnem mestu, ki pripada galoitalskim točkam Aidone, Sperlinga in San Fratello v sistemu mreže AIS; 27-54
181. Gerhard ERNST: Textes privés des XVIIe et XVIIIe siècles: problèmes et chances d'une édition - Zasebna besedila 17. in 18. stoletja: problemi in možnosti njihove izdaje; 55-68
182. Jacqueline OVEN: Traduction de la particule de mise en relief slovène tudi en français: cas de disparitions surprenantes - Prevajanje slovenskega členka tudi v francoščino: presenetljiva izginotja v francoskem prevodu; 69-80
183. Ana María SERRADILLA CASTAÑO: Abrir, cerrar, subir y bajar: la productividad de los verbos de movimiento como elementos constitutivos de locuciones idiomáticas en español – Abrir, cerrar, subir y bajar: produktivnost glagolov gibanja kot sestavnih elementov idiomatičnih izrazov v španskem jeziku; 81-100
184. Pierre SWIGGERS: Una aportación flamenca a la gramaticografía del español: la »Spaensen grammatica« de Arnaldo de la Porte (1659) – Flamski prispevek

- k španskemu slovniciarstvu: »Spaensen grammatica« Arnalda de la Portreja (1659); 101-112
185. María Pilar SANCHÍS SERDAN: El Diálogo de la lengua de Juan de Valdés: modelo humanista de defensa de la lengua castellana en el siglo XVI - Juan de Valdés, Diálogo de la lengua: humanistični model obrambe kastiljskega jezika v šestnajstem stoletju; 113-125
186. Dieter MESSNER: Sobre traducciones fracasadas - O neuspelih prevodih; 127-136
187. Lucija ČOK: Lingue e culture nel dibattito sulle identità europee - Jeziki in kulture v diskurzu o evropskih identitetah; 137-142
188. Rienzo PELLEGRINI: Mezzogiorno alpino di Giosuè Carducci - Mezzogiorno alpino Giosuèja Carduccija; 143-156
189. Emanuele BANFI: La Grecia moderna e la formazione della Néa Ellinikí Koiní (NEK) - Moderna Grčija in nastanek Néa Ellinikí Koiní (NEK); 157-168
190. Jerneja KAVČIČ: Syntactic variants and natural syntax: Greek infinitive clauses dependent on verbs of speaking and thinking - Skladenjske različice in naravna skladnja: grški nedoločniški stavki, odvisni od glagolov govorenja in mišljenja; 169-177
191. Frančiška TROBEVŠEK DROBNAK: Distribution of infinitive markers in Chaucer's Canterbury Tales - Distribucija označevalcev nedoločniškega obrazila v Chaucerjevih Canterburyjskih zgodbah; 179-195
192. Teodor PETRIČ: Inkohärente Infinitivkonstruktionen in deutschen Untertiteln - Inkohärentni nedoločniški polstavki v podnapisih; 197-208
193. Urška VALENČIČ ARH, Tjaša JAKOP: Semantik und Motivation in der Entstehung der Phraseme mit der Komponente *denar* (dt. *Geld*) - Semantika in motivacija frazemov s komponento *denar*; 209-223
194. Sašo ŽIVANOVIĆ: Conjunctive and prepositional comparatives in Slovenian - Vežniške in predložne primerniške zgradbe v slovenščini; 225-240
195. Kenneth SHIELDS: More on the origin of the Indo-European feminine gender: a reply to Ledo-Lemos - Še nekaj o izvoru indoevropskega ženskega spola: odgovor Ledo-Lemosu; 241-248
196. Simona KLEMENČIČ: On the methodology of dialectology of language families: determining the existence of a dialect areal - K metodologiji dialektologije jezikovnih družin: sklepanje o obstoju narečnega areala; 249-267

## LI, 2011

197. Michel ROCHÉ: Pression lexicale et contraintes phonologiques dans la dérivation en *-aie* du français - Leksikalni vpliv in fonološke omejitve v izpeljavi s pripono *-aie* v francoščini; 5-22



198. Irena STRAMLJIČ BREZNIK, Ines VORŠIČ: Word-formational productivity of the Slovene language in the case of sports neologisms - Besedotvorna produktivnost slovenščine na primeru novotvorjenk s področja športa; 23-38
199. Alexandra BAGASHEVA: Beyond dichotomies: on the nature and classification of compound verbs in English - Onstran dihotomij: o naravi in tipologiji glagolskih zloženek v angleščini; 39-63
200. Maris CAMILLERI: Island morphology: morphology's interactions in the study of stem patterns - Otoška morfologija: vloga morfoloških interakcij pri raziskovanju vzorcev osnove; 65-85
201. Arnaud LÉTURGIE: Un cas d'extragrammaticalité particulier: les amalgames lexicaux fantaisistes - Poseben primer izvensistemskosti: izmišljeni leksikalni amalgami; 87-104
202. Silvia CACCHIANI: On unfamiliar Italian lexical blends from names and nouns - O manj pogosti vrsti leksikalnih spojenk iz imen in samostalnikov; 105-120
203. Thomas SCHWAIGER: The relation between prototypical and marginal morphology: the case of reduplicative constructions - Razmerje med prototipsko in marginalno morfologijo: primer reduplikativnih konstrukcij; 121-134
204. Eva SICHERL, Andreja ŽELE: Nominal diminutives in Slovene and English - Samostalniške manjšalnice v slovenščini in angleščini; 135-142
205. Marie-Anne BERRON, Marie MOUTON: La morphologie marginale dans le slam: comparaison interlinguale France vs Allemagne - Marginalna morfologija v slamu: francosko-nemška medjezikovna primerjava; 143-155
206. Géraldine WALTHER: Measuring morphological canonicity – Merjenje morfološke kanoničnosti; 157-180
207. Tatjana MARVIN : Phases at the word level - Faze na besednem nivoju; 181-199
208. Patrizia CORDIN: From verbal prefixes to direction/result markers in Romance - Od glagolskih predpon do označevalcev smeri/rezultata v romanskih jezikih; 201-216
209. Erich R. ROUND: Morphemes as a level of representation capture unity of exponence across the inflection–derivation divide - Morfomi kot predstavitevna ravnina zajamejo enotnost izraznih podob ne glede na razkol med oblikoslovljem in besedotvorjem; 217-230
210. Georgia ZELLOU: Moroccan Arabic borrowed circumfix from Berber: investigating morphological categories in a language contact situation - Berberski circumfiks v maroški arabščini: raziskovanje morfoloških kategorij na primeru jezikov v stiku; 231-244
211. Chikako SHIGEMORI BUČAR: Creative competence in borrowings: words of Japanese origin in Slovene - Ustvarjalna kompetenca v izposojanju: besede japonskega izvora v slovenščini; 245-262

212. Varja CVETKO-OREŠNIK: Naturalness: some synchronic data of old Indian - Naravno jezikoslovje: nekaj sinhronih podatkov iz stare indijščine; 263-272
213. Douglas LIGHTFOOT: Testing the suffixoidization of German *-mann* ‘-man’ - Preverjanje sufiksoidizacije nemškega »-mann« ‘moški’; 273-283
214. Javier E. DÍAZ VERA: Derivation in a word-based morphology: on the origin of old English verbs of perception, cognition and emotion - Izpeljevanje v besednoutemeljeni morfologiji: o izvoru staroangleških glagolov zaznave, spoznavanja in čustev; 285-290
215. Metka FURLAN: The archaic word-formational relationship between the Proto-slavic noun *\*pol'e* and the Hittite adjective *palhi-* - Arhaično besedotvorno razmerje med praslovanskim samostalnikom *\*pol'e* in hetitskim pridevnikom *palhi-*; 291-297
216. Marc PLÉNAT, Camille PLÉNAT: La liaison de l'adjectif sur le nom en français: morphologie, syntaxe, phonologie - Vežanje pridevnika na samostalnik v francoščini: morfologija, sintaksa, fonologija; 299-315
217. Janez OREŠNIK: The imperative in natural syntax - Velelnik v naravni skladnji; 317-332
218. Mojca SCHLAMBERGER BREZAR: Le gérondif et le participe présent et leur évolution vers la grammaticalisation: étude contrastive du slovène et du français - Deležja in deležniki sedanjega časa in njihov razvoj v smeri gramatikalizacije: slovensko-francoska kontrastivna študija; 333-348
219. Gašper ILC: Jespersen's cycle in Slovenian - Jespersenov krog v slovenščini; 349-363
220. Mojca SMOLEJ: Qu'est-ce que le slovène et le français ont en commun ? Les articles défini et indéfini - Kaj imata slovenščina in francoščina skupnega? Določni in nedoločni člen; 365-375
221. Barbara Sandig: *Textstilistik des Deutschen*. Berlin New York: Walter de Gruyter, 2006 (Stojan BRAČIČ); 377-382
222. Toshiko Yamaguchi: *Japanese linguistics: an introduction*. London/New York, 2007 (MAGNÚS PÉTURSSON); 383-384
223. Paul Videsott: *Padania scrittologica. Analisi scrittologiche e scrittometriche di testi in italiano settentrionale antico dalle origini al 1525*. Beihefte zur Zeitschrift für romanische Philologie, 343. Tübingen: Max Niemeyer Verlag, 2009 (Ludwig FESENMEIER); 385-389
224. ... *Sand carried by a stream... Scritti in onore di Vincenzo Orioles*, a cura di Raffaella Bombi e Fabiana Fusco. Udine: ed. Forum, 2009 (Paolo DI GIOVINE); 390-394
225. Jan Radimský: *Les composés italiens actuels*. Paris: Cellule de recherche en linguistique, 2006. (Gregor PERKO); 395-398
226. Dany Amiot et al. : *Ars grammatica: Hommages à Nelly Flaux*. Bern: Peter Lang, 2011. (Gregor PERKO); 399-402
227. Metka ŠORLI: *Linguistica XLI-LI*; 403-442

**INDEX DES AUTEURS DES CONTRIBUTIONS (noms en majuscules)**  
**INDEX DES AUTEURS CITÉS DANS LES TITRES DES CONTRIBUTIONS**  
**(noms en minuscules)**

**SEZNAM AVTORJEV PRISPEVKOV (priimki z velikimi črkami)**  
**SEZNAM AVTORJEV, CITIRANIH V NASLOVIH PRISPEVKOV**  
**(priimki z malimi črkami)**

Dany Amiot 226  
Marcello Aprile 117  
Sandra AUGENDRE 169  
Matjaž BABIČ 12, 19, 64, 90  
Alexandra BAGASHEVA 199  
Emmanuele BANFI 12, 189  
Giorgio Bassani 42  
Jan Niecisław Baudouin de Courtenay 178  
Roland BAUER 179  
Roland Bauer (ur. / sous la direction de) 49  
Andrej BEKEŠ 158  
Walter Belardi 67  
Marie-Anne BERRON 205  
Gaetano BERRUTO 161  
David BIZJAK 39, 93  
Adam Bohorič 83  
Raffaella Bombi (ur. / sous la direction de) 224  
Giuliano BONFANTE 1  
Ignacio Bosque (ur. / sous la direction de) 13  
Stojan BRAČIČ 34, 35, 155, 221  
Maria Bruniera 134  
Laura BURAZER 98  
Concepción CABRILLANA 122  
Silvia CACCHIANI 202  
Maris CAMILLERI 200  
Giosué Carducci 188  
Augusto CARLI 162  
Louis-Ferdinand Céline 141  
Maria Rosaria CERASUOLO PERTUSI 20

Miguel de Cervantes Saavedra 151  
Geoffrey Chaucer 191  
Jean de Clercq (ur. / sous la direction de) 63  
Patrizia CORDIN 208  
Michele A. CORTELAZZO 166  
Rada Cossutta 132  
Franco Crevatin 132  
Varja CVETKO-OREŠNIK 10, 102, 212  
Lucija ČOK 187  
Željka Čorak 7  
Darko ČUDEN 134  
Jurij Dalmatin 105  
Roberto DAPIT 177  
Arnaldo De la Porte 184  
Violeta Demonte (ur. / sous la direction de) 13  
Aleksandra DERGANČ 108  
Cristiana De Santis 119  
Paola DESIDERI 163  
Javier E. DÍAZ VERA 214  
Paolo DI GIOVINE 224  
Wolfgang U. DRESSLER 74  
Gerhard ERNST 181  
Alja FERME 106  
Alonso Fernández de Avellaneda 151  
Ludwig FESENMEIER 223  
Goran Filipi 46  
Nelly Flaux 226  
Marco Forni 67  
Metka FURLAN 159, 215  
Fabiana FUSCO 171  
Fabiana Fusco (ur. / sous la direction de) 224  
Florence GACOIN-MARKS 85  
Marek GAWELKO 96  
Renato GENDRE 16, 38  
Jules Gilliéron 62  
Hans GOEBL 95, 180

Hans Goebel (ur. / sous la direction de) 49  
Marija GOLDEN 71  
Gaston Gross 119  
Ada GRUNTAR JERMOL 129  
Roberto GUSMANI 175  
Claude Hagège 114  
Eric P. HAMP 15, 51  
József Herman 11  
Louis Hjelmslev 75  
Matej HRIBERŠEK 37  
Gabriele Iannàccaro 115  
Gašper ILC 29, 80, 219  
Roxana IORDACHE 18, 89  
Roxana Iordache 50, 65, 66  
Tjaša JAKOP 193  
Nina Janich 35  
Josip JERNEJ 53  
Vida JESENŠEK 30  
Otto Jespersen 219  
Juan Ramón Jiménez 150  
Ana JUVANČIČ MEHLE 73  
Branka KALENIĆ RAMŠAK 151  
Jerneja KAVČIČ 5, 56, 88, 190  
Jana KENDA 146  
Simona KLEMENČIČ 196  
Smilja KOMAR 153  
Snježana Kordić 48  
Erika KRŽIŠNIK 84  
Meta LAH 58, 125, 140  
Peter Lauwers (ur. / sous la direction de) 62  
Francisco José Ledo-Lemos 195  
Arnaud LÉTURGIE 201  
Vladimir Levstik 85  
Douglas LIGHTFOOT 213  
Nico Lioce (ur. / sous la direction de) 63  
MAGNÚS PÉTURSSON 52, 77, 128, 164, 222

Johannes Malalas 5  
Witold MAŃCZAK 11, 81  
Thomas L. MARKEY 104  
Jasmina MARKIČ 13, 111, 149  
Tatjana MARVIN 45, 79, 207  
Ranko Matasović 31  
Pavle Merku 14  
Darja MERTELJ 145  
Dieter MESSNER 186  
Tjaša MIKLIČ 21, 42, 91, 123, 144  
Milena MILOJEVIĆ SHEPPARD 29, 80  
Marie MOUTON 205  
Žarko MULJAČIĆ 87  
Žarko Muljačić 47  
Eva NÚÑEZ 124  
Irena OREL 105  
Silvana OREL KOS 154  
Janez OREŠNIK 2, 28, 44, 121, 138, 165, 217  
Janez Orešnik 71, 73  
Vincenzo ORIOLES 43, 172  
Vincenzo Orioles 224  
Jacqueline OVEN 57, 182  
Martina OŽBOT 3, 55, 99, 123, 135, 148, 160  
Sorin PALIGA 17  
Giuseppe Patota 68, 118  
Rienzo PELLEGRINI 188  
Gregor PERKO 4, 32, 59, 141, 225, 226  
Teodor PETRIČ 101, 192  
Jacqueline Picoche 32  
Barbara PIHLER 150  
Titus Maccius Plautus 19, 90  
Camille PLÉNAT 216  
Marc PLÉNAT 216  
Vladimir POGAČNIK 23, 63, 112  
Heinz Dieter POHL 82  
Janja POLAJNAR 156

Michele Prandi 119  
Marjeta PRELESNIK DROZG 137  
Mirjam PREMRL 170  
Maurizio Puntin 61  
Raymond Queneau 143  
Jan Radimský 225  
Michel ROCHÉ 197  
Jean-Claude Rolland 32  
Erich R. ROUND 209  
Leyre RUIZ DE ZAROBÉ 6  
Oana Sălișteanu Cristea 33  
María Pilar SANCHÍS CERDAN 185  
Barbara Sandig 221  
Mojca SCHLAMBERGER BREZAR 24, 218  
Thomas SCHWAIGER 203  
Ana María SERRADILLA CASTAÑO 183  
Kenneth SHIELDS 76, 195  
Chikako SHIGEMORI BUČAR 107, 211  
Eva SICHERL 109, 157, 204  
Marie-Rose Simoni-Aurembou (ur. / sous la direction de) 62  
Mitja SKUBIC 14, 40, 46, 47, 61, 62, 70, 94, 113, 131, 132, 133  
Mitja Skubic 135, 137  
Gunver SKYTTE 167  
Mojca SMOLEJ 127, 220  
Rosanna SORNICOLA 168  
Liliana SPINOZZI MONAI 178  
Irena STRAMLJIČ BREZNIK 198  
Pierre SWIGGERS 75, 120, 184  
Pierre Swiggers (ur. / sous la direction de) 62, 63  
Maja ŠABEC 152  
Agata ŠEGA 147  
Sergij Šlenc 131  
Metka ŠORLI 227  
Pavao TEKAVČIĆ 7, 33, 49, 50, 65, 66, 67, 68, 69, 86, 114, 115, 116, 117, 118, 119  
Pavao Tekavčić 113  
Publius Terentius Afer 19

Karmen TERŽAN-KOPECKY 100  
Zuzanna TOPOLIŃSKA 97  
Jože TOPORIŠIČ 83  
Fiorenzo TOSO 41, 92, 173  
Ivan Trinko 133  
Frančiška TROBEVŠEK DROBNAK 191  
Juan de Valdés 185  
Urška VALENČIČ ARH 193  
Sonia VAUPOT 27, 143  
Federico VICARIO 36, 176  
Paul VIDESOTT 9  
Paul Videsott 223  
Claude Vincenot 34  
Primož VITEZ 25, 142  
Ines VORŠIČ 198  
Alenka VRBINC 60, 110, 130  
Marjeta VRBINC 60, 110, 130  
Zorica VUČETIĆ 8, 22, 54, 103  
Géraldine WALTHER 206  
John WEINSTOCK 78  
Roger Wright 11  
Toshiko YAMAGUCHI 77  
Toshiko Yamaguchi 222  
Alberto ZAMBONI 174  
Georgia ZELLOU 210  
Marina ZORMAN 31, 48  
Ana ZWITTER VITEZ 26, 139  
Andreja ŽELE 109, 157, 204  
Sašo ŽIVANOVIĆ 194  
Nataša ŽUGELJ 126



## INDEX THÉMATIQUE - STVARNO KAZALO

*Les chiffres se rapportent aux numéros des contributions de La table chronologique*

- accentuation 83, 207
- adjectif 215, 216
- adverbe 92, 158
- albanais 15, 51
- allemand 30, 35, 100, 129, 155, 156, 157, 179, 186, 192, 205, 213, 221
  - ~ autrichien 186
- amalgames lexicaux 201, 202
- anglais 28, 44, 60, 79, 109, 110, 121, 130, 153, 154, 199, 204, 207
  - ancien ~ 214
  - moyen ~ 191
- anthroponimie 176
- anticipation 91, 144
- antonymes 36
- aoriste 102
- arabe marocain 210
- article 27, 220
- article de journal 77
- aspect 21, 111
- atlas linguistique 9, 46, 95, 132, 180
- bibliographie 73, 137
- bilinguisme 124
- biologie du langage 62
- calques 178
- castillan 38, 185
- celtique 1, 104
- codes grammaire 110
- cohérence 3, 167
- collocations 105, 109, 129, 157
- commentaire journalistique 30
- communication interculturelle 99, 100, 187
- complément hypothétique 18
- concession 50
- connecteurs 24
- connexion implicite 167
- construction détachée
- naglaševanje
- pridevnik
- prislov
- albanščina
- nemščina
  - avstrijska ~
- spojenke
- angleščina
  - stara ~
  - srednja ~
- osebna imena
- pogled naprej
- antonimi
- aorist
- maroška arabščina
- člen
- časopisni članek
- glagolski vid
- jezikovni atlas
- bibliografija
- dvojezičnost
- jezikovna biologija
- kalki
- kastiljščina
- keltščina
- slovnične kode
- koherenca
- kolokacije
- časopisni komentar
- medkulturno sporazumevanje
- dopolnilo pogojnosti
- dopustnost
- povezovalci
- implicitna zveza
- polstavek

- ~ infinitive 56, 190, 191, 192
- ~ participiale 45
- constructions causatives 107
- constructions comparatives
  - ~ conjonctionnelles 194
  - ~ prépositionnelles 194
- constructions réduplicatives 203
- contexte
  - ~ linguistique 169
  - ~ situationnel 98
- corse 41
- croate 7, 8, 48
  - ~ de Molise 87
- cycle de Jespersen 219
- dalmate 47
- danois 167
- datif étique 94
- déclinaison 44, 76, 159
- déplacement d'un verbe 29
- détermination 122
  - ~ nominale 27
- dialecte
  - ~ allemand en Italie (Sappada) 134
  - ~ d'Arenzano 173
  - ~ s de Corse 92
  - ~ s gallo-italiques 180
  - ~ s siciliens 180
  - ~ s slovènes 132, 147, 178
    - ~ du Val de la Tour 177
    - ~ resian 177
- dialectologie 92, 95, 132, 134, 180, 196, 223
- dialectométrie 95, 179, 180
- dictionnaires 116
  - ~ bilingues 59, 131
  - ~ monolingues 4, 32, 110, 130
  - emploi des ~ 60, 130
- didactique des langues
  - 6, 55, 58, 60, 63, 123, 125, 126, 128, 130, 144, 145, 170
- didascalie 142
- discours
  - analyse du ~ 146, 149, 150
- nedoločniški ~
- deležniški ~
- vzročniške zgradbe
- primerniške zgradbe
  - vezniške ~
  - predložne ~
- podvojitvene zgradbe
- kontekst
  - jezikovni ~
  - situacijski ~
- korziščina
- hrvaščina
  - moliška ~
- Jespersenov krog
- dalmatščina
- danščina
- etični dativ
- sklanjatev
- glagolski premik
- določnost
  - samostalniška ~
- narečje
  - nemško ~ v Italiji (Sappada)
  - ~ Arenzana
  - korziška ~a
  - galoitalska ~a
  - sicilijanska ~a
  - slovenska ~a
    - tersko ~
    - rezijsko ~
- dialektologija
- dialektometrija
- slovarji
  - dvojezični ~
  - enojezični ~
  - uporaba ~ev
- didaktika (tujega) jezika
- didaskalije
- govor / diskurz
  - analiza diskurza

~ indirect libre 42, 144  
 ~ poétique 15  
 ~ rapporté 6  
 ~ spontané 26, 139  
 disparition des langues 120  
 e muet 79  
 emprunts 8, 9, 43, 82, 116, 148, 211  
 enclitiques 84  
 équivalents 59  
 espagnol  
     13, 38, 111, 150, 183, 184, 185, 186  
     ~ d'Amérique latine 149  
 étrusque 104  
 étymologie  
     15, 17, 20, 38, 51, 112,  
     117, 173, 174, 175  
     ~ populaire 86  
 expression  
     ~ écrite 125, 126  
     ~ parlée 25, 26, 139, 146  
 expressions idiomatiques 4, 84, 183, 193  
 fait divers 140  
 famille de langues 196  
 faux amis 108  
 flamand 184  
 formation des mots  
     8, 22, 74, 79, 103, 157, 172,  
     197, 198, 199, 202, 204, 207,  
     209, 210, 213, 214, 215  
 formes verbales 21, 123, 144, 149, 170  
 français  
     4, 6, 23, 24, 25, 26, 27, 32, 39, 57,  
     58, 59, 63, 85, 95, 125, 126, 127, 138,  
     139, 143, 179, 181, 182, 197, 201,  
     205, 216, 218, 220  
 frioulan  
     39, 40, 70, 93, 94, 175, 176, 177, 178  
 génitif 44  
 genre 171, 211  
     ~ féminin 195  
 géographie linguistique 9, 62, 132, 180  
 gérondif 218

polpremi govor  
 pesniški diskurz  
 poročani govor  
 spontani ~  
 izumiranje jezikov  
 polglasnik  
 izposojenke  
 klitike  
 prevodne ustreznice  
 španščina  
  
     ameriška ~  
 etruščina  
 etimologija  
  
     ljudska ~  
 izražanje  
     pisno ~  
     govorno ~  
 idiomi  
 črna kronika  
 jezikovna družina  
 lažni prijatelji  
 flamščina  
 besedotvorje  
  
 glagolske oblike  
 francoščina  
  
 furlanščina  
  
 rodilnik  
 spol  
     ženski ~  
 zemljepisno jezikoslovje  
 deležje

- gotique 16  
grammaire 63, 68, 130  
~ comparée 34  
~ descriptive 13  
~ historique 168, 184  
grammaticalisation 168, 218  
grec 5, 12, 16, 56, 64, 88, 190  
~ moderne 189  
hitite 215  
identité  
~s européennes 187  
~s linguistiques 187  
imparfait 21, 23, 37  
impératif 217  
index 227  
infinitif 56, 191  
intonation 26, 153  
islandais 52, 77, 128  
istroroumain 46  
italianismes  
~ en croate 8  
~ en ladin 9  
~ en slovène 148  
italien  
3, 21, 22, 36, 42, 43, 53, 54, 55, 68, 91,  
103, 117, 118, 119, 123, 131, 144, 145,  
146, 166, 167, 168, 169, 170, 171, 172,  
173, 179, 202, 223, 225  
japonais 107, 158, 211, 222  
kayardild (langue australienne) 209  
ladin 9, 67, 179  
langue (langage) de spécialité  
langage de l'art 103  
~ juridique 22, 129  
~ politique 54, 163  
~ publicitaire 35, 156  
langues 164  
~ anatoliennes 159  
~ baltiques 159  
~ indo-européennes 10, 76, 96, 159, 195  
~ romanes 65, 114, 116, 208  
~ slaves 17, 159  
gotščina  
slovnica  
primerjalna ~  
opisna ~  
historična ~  
gramatikalizacija  
grščina  
moderna ~  
hetitščina  
identiteta  
evropske ~e  
jezikovne ~e  
nedovršnik  
velelnik  
kazala  
nedoločnik  
intonacija  
islandščina  
istroromunščina  
italianizmi  
~ v hrvaščini  
~ v ladinščini  
~ v slovenščini  
italijanščina  
japonščina  
kayardild (avstralski jezik)  
ladinščina  
strokovni jezik  
jezik umetnosti  
pravniški ~  
politični ~  
reklamni ~  
jeziki  
anatolski ~  
baltski ~  
indoevropski ~  
romanski ~  
slovanski ~

- langues en contact  
     8, 9, 43, 82, 116, 147, 148,  
     177, 178, 179, 180, 210, 211  
 langues ergatives 165  
 latin  
     11, 18, 19, 37, 50, 64, 65, 66, 89,  
     90, 114, 122, 168  
 latinismes dans les langues romanes 116  
 lexicographie 4, 32, 59, 110, 116, 131, 171  
 lexicologie  
     8, 9, 17, 20, 22, 38, 43, 54, 108,  
     116, 148, 173, 174  
 liaison 138, 216  
 linguistes  
     ~ croates 113  
     ~ slovènes 71, 73, 135, 137  
 linguistique comparée  
     15, 16, 17, 26, 29, 31, 34, 38, 39, 45,  
     51, 57, 80, 81, 86, 96, 104, 107, 109,  
     127, 139, 154, 157, 159, 165, 167,  
     195, 196, 204, 205, 215, 218, 220, 226  
 linguistique générale 75  
 linguistique textuelle  
     3, 24, 35, 140, 144,  
     146, 149, 150, 167, 221  
 littérature  
     ~ anglaise 191  
     ~ autrichienne 186  
     ~ croate 7  
     ~ espagnole 150, 151, 152, 186  
     ~ française 141, 143  
     ~ italienne 42, 188  
     ~ ladine 67  
     ~ slovène 57, 182  
 locatif 76  
 macédonien 97  
 marqueurs  
     ~ de direction 208  
     ~ de l'infinif 191  
     ~ de négation 219  
     ~ discursifs 146  
 métaphore 78
- jeziki v stiku  
  
 ergativni jeziki  
 latinščina  
  
 latinizmi v romanskih jezikih  
 slovaroslovje  
 besedoslovje  
  
 vezanje  
 jezikoslovci  
     hrvaški ~  
     slovenski ~  
 primerjalno jezikoslovje  
  
 splošno jezikoslovje  
 besediloslovje  
  
 književnost  
     angleška ~  
     avstrijska ~  
     hrvaška ~  
     španska ~  
     francoska ~  
     italijanska ~  
     ladinska ~  
     slovenska ~  
 mestnik  
 makedonščina  
 označevalci  
     ~ smeri  
     nedoločniški ~  
     ~ zanikanja  
     diskurzni ~  
 metafora

modalité déontique 23  
 modèle morphologique »parsli« 206  
 monolinguisme 124  
 morphèmes  
   ~ dérivationnels 109, 157, 210, 213  
   ~ libres 208  
 morphologie 159, 168, 200, 206, 209,  
   210, 211, 216, 218, 226  
   ~ canonique 206  
   ~ marginale 201, 202, 203, 205  
   ~ naturelle 74  
   ~ prototypique 203  
 morphomes 209  
 morphosyntaxe 147  
   ~ naturelle 2, 28, 44, 102, 212  
 mots composés 74, 199, 225  
 mots dérivés 79, 197, 204, 207, 214  
 mots valises – voir amalgames lexicaux  
 négation 80, 219  
 néologismes 198  
 niveaux de langue 143  
 nom 215, 216  
   ~s diminutifs 204  
 nombre 211  
 norme (linguistique) 25, 189  
 ordre de mots 36, 90, 122  
 parfait 77, 102  
 parler  
   ~ de Celje 106  
   ~ de Trentino 208  
   ~s berbères 210  
   ~ triestin 20  
 participe 5  
   ~ passé 45  
   ~ présent 218  
 particule 127  
   ~ aussi 57, 182  
   ~ »si« 146  
 pédagogie interculturelle 187  
 périphrases verbales 39, 111, 149  
 phonologie 79, 106, 207, 216  
 phrase

deontična modalnost  
 oblikoslovni model »parsli«  
 enojezičnost  
 morfemi  
   besedotvorni ~  
   prosti ~  
 oblikoslovje  
   ~vna kanoničnost  
   marginalno ~  
   naravno ~  
   prototipsko ~  
 morfomi  
 oblikoskladnja  
   naravna ~  
 zloženke  
 izpeljanke  
 spojenke  
 zanikanje  
 novotvorjenke  
 jezikovne ravni  
 samostalnik  
   ~ške manjšalnice  
 število  
 jezikovna norma  
 besedni red  
 dovršnik  
 govor  
   celjski ~  
   trentinski ~  
   berberski ~i  
   triestinski ~  
 deležnik  
   pretekli ~  
   sedanji ~  
 členek  
   ~ tudi  
   ~ »si«  
 medkulturna pedagogika  
 glagolske perifraze  
 fonologija  
 poved

~ impérative 145  
 ~ simple 53  
 phrasèmes 4, 84, 183, 193  
 pléonasme 88  
 plurilinguisme 187  
 préposition 105, 109  
 pronom 44, 92  
   ~ démonstratif »ipse« 168  
   ~ négatif 80  
   ~ personnel 19  
   ~ relatif 40  
 prononciation 128, 138  
 proposition  
   ~ interrogative 29, 121  
   ~ négative 219  
 prosodie de l'énonciation 139  
 proto-slave 215  
 quantificateur 27  
 recensions  
   12, 13, 14, 31, 32, 33, 34, 35, 46,  
   47, 48, 49, 50, 61, 62, 63, 64, 65,  
   66, 67, 68, 69, 70, 115, 116, 117,  
   118, 119, 131, 132, 133, 134, 221,  
   222, 223, 224, 225, 226  
 référence 98  
 rétoroman 179  
 rétrospection 144  
 roumain 17  
 russe 108  
 russismes en italien 43  
 same 78  
 sémantique 48, 177, 193  
   ~ textuelle 155  
 slam 205  
 slave commun – voir proto-slave  
 slovène  
   3, 26, 51, 55, 57, 59, 80, 82, 83, 84,  
   85, 100, 105, 107, 108, 109, 112, 127,  
   129, 131, 139, 147, 148, 154, 157, 177,  
   178, 182, 193, 194, 198, 204, 211,  
   218, 219, 220  
 sociolinguistique 120, 171  
 velelna ~  
 enostavčna ~  
 frazemi  
 pleonazem  
 večjezičnost  
 predlog  
 zaimek  
   kazalni ~ »ipse«  
   nikalni ~  
   osebni ~  
   oziralni ~  
 izgovorjava  
 stavek  
   vprašalni ~  
   nikalni ~  
 prozodija izrekanja  
 praslovanščina  
 količinski določevalnik  
 recenzije  
 nanašanje  
 retoromanščina  
 pogled nazaj  
 romunščina  
 ruščina  
 rusizmi v italijanščini  
 samijski jezik  
 pomenoslovje  
   besedilno ~  
 slam  
 praslovanščina  
 slovenščina  
 sociolingvistika

sonantes en final de mot 106  
 sons 52, 106  
 sous-titrage 192  
 stéréotypes 171  
 stylistique 16, 154  
 subordonnée de condition 89  
 syllabation 79  
 syntagme 93  
   ~ nominal générique 97  
 syntaxe  
   18, 29, 45, 48, 50, 53, 64, 66, 80, 89,  
   119, 122, 194, 216, 217, 219, 226  
   ~ naturelle  
   56, 121, 138, 165, 190, 212  
 »tabula gratulatoria« 72, 136  
 temps verbaux  
   21, 23, 37, 77, 123, 144, 149, 154  
 terminologie maritime 8  
 texte  
   authenticité de ~ 58  
   genres de ~ 24, 58, 140  
   ~ scénique 142  
   ~s argumentatifs 24  
   types de ~ 30, 35, 140  
 théorie du langage 161  
 théorie du langage naturel  
   2, 10, 28, 44, 56, 74, 88, 100, 101,  
   102, 121, 138, 165, 190, 212  
 toponymie 14, 61, 82, 112, 175, 177  
 traduction 55, 100, 108, 129, 192  
   ~ littéraire 3, 85, 105, 182, 186  
 traductologie 99  
 transitivité 165  
 valence 109, 157  
 variation (linguistique) 49  
   ~ intralinguistique 161  
 verbe  
   21, 23, 29, 36, 37, 39, 56, 77, 84, 123,  
   144, 149, 154, 157, 169, 170, 208, 212  
   ~s auxiliaires 111, 166  
   ~s causatifs 214  
   ~s déclaratifs 88, 105

zvočniško izglasje  
 glasovi  
 podnaslavljanje  
 stereotipi  
 stilistika  
 pogojni odvisnik  
 zlogovanje  
 sintagma  
   generična imenska ~  
 skladnja  
  
   naravna ~  
  
 »tabula gratulatoria«  
 glagolski časi  
  
 pomorska terminologija  
 besedilo  
   avtentičnost ~a  
   tipi ~a  
   scensko besedilo  
   utemeljevalna besedila  
   besedilne vrste  
 teorija jezika  
 teorija jezikovne naravnosti  
  
 krajevna imena  
 prevajanje  
   literarno ~  
 prevodoslovje  
 prehodnost  
 vezljivost  
 jezikovna raznolikost  
 znotrajjezikovna ~  
 glagol  
  
   pomožni ~i  
   kavzativni ~ i  
   ~i govorjenja



~s de cognition 214  
~s de déplacement 111, 183  
~s d'émotion 214  
~s de perception 214  
~s modaux 23, 166  
vieil indien 102, 212  
vieux slave 81  
vieux slave alpin 82  
vitalité des langues 162

~i spoznanja  
~i gibanja  
~i čustev  
~i zaznave  
modalni ~i  
stara indijščina  
stara cerkvena slovanščina  
karantanščina  
jezikovna vitalnost

## STVARNO KAZALO – INDEX THÉMATIQUE

*Številke se nanašajo na prispevke, oštevilčene v Kronološkem pregledu*

- albanščina 15, 51
- alpska slovanščina - gl. katantanščina
- analiza diskurza 146, 149, 150
- angleščina 28, 44, 60, 79, 109, 110, 121, 130, 153, 154, 199, 204, 207
  - srednja ~ 191
  - stara ~ 214
- antonimi 36
- aorist 102
- besedilne vrste 30, 35, 140
- besedilo
  - avtentičnost ~a 58
  - scensko ~ 142
  - tipi ~a 24, 58, 140
  - utemeljevalna ~a 24
- besediloslovje 3, 24, 35, 140, 144, 146, 149, 150, 167, 221
- besedni red 36, 90, 122
- besedoslovje 8, 9, 17, 20, 22, 38, 43, 54, 108, 116, 148, 173, 174
- besedotvorje 8, 22, 74, 79, 103, 157, 172, 197, 198, 199, 202, 204, 207, 209, 210, 213, 214, 215
- bibliografija 73, 137
- časopisni članek 77
- časopisni komentar 30
- člen 27, 220
- členek 127
  - ~ tudi 57, 182
  - ~ »si« 146
- črna kronika 140
- dalmatščina 47
- danščina 167
- deležje 218
- deležnik 5
  - pretekli ~ 45
  - sedanji ~ 218
- albanais
- vieux slave alpin
- analyse du discours
- anglais
  - moyen ~
  - ancien ~
- antonymes
- aoriste
- types de textes
- texte
  - authenticité de ~
  - ~ scénique
  - genres de ~
  - ~s argumentatifs
- linguistique textuelle
- ordre de mots
- lexicologie
- formation des mots
- bibliographie
- article de journal
- commentaire journalistique
- article
- particule
  - ~ aussi
  - ~ »si«
- fait divers
- dalmate
- danois
- gérondif
- participe
  - ~ passé
  - ~ présent

deontična modalnost 23	modalité déontique
dialektologija 92, 95, 132, 134, 180, 196, 223	dialectologie
dialektometrija 95, 179, 180	dialectométrie
didaktika (tujega) jezika 6, 55, 58, 60, 63, 123, 125, 126, 128, 130, 144, 145, 170	didactique des langues
didaskalije 142	didascalie
določnost 122	détermination
samostalniška ~ 27	~ nominale
dopolnilo pogojnosti 18	complément hypothétique
dopustnost 50	concession
dovršnik 77, 102	parfait
dvojezičnost 124	bilinguisme
enojezičnost 124	monolinguisme
ergativni jeziki 165	langues ergatives
etični dativ 94	datif étique
etimologija 15, 17, 20, 38, 51, 112, 117, 173, 174, 175	étymologie
ljudska ~ 86	~ populaire
etruščina 104	étrusque
flamščina 184	flamand
fonologija 79, 106, 207, 216	phonologie
francoščina 4, 6, 23, 24, 25, 26, 27, 32, 39, 57, 58, 59, 63, 85, 95, 125, 126, 127, 138, 139, 143, 179, 181, 182, 197, 201, 205, 216, 218, 220	français
frazemi 4, 84, 183, 193	phrasèmes
furlanščina 39, 40, 70, 93, 94, 175, 176, 177, 178	frioulan
glagol 21, 23, 29, 36, 37, 39, 56, 77, 84, 123, 144, 149, 154, 157, 169, 170, 208, 212	verbe
~i čustev 214	~s d'émotion
~i gibanja 111, 183	~s de déplacement
~i govorjenja 88, 105	~s déclaratifs
~i spoznanja 214	~s de cognition
~i zaznave 214	~s de perception
kavzativni ~i 214	~s causatifs
modalni ~i 23, 166	~s modaux
pomožni ~i 111, 166	~s auxiliaires
glagolske oblike 21, 123, 144, 149, 170	formes verbales

glagolske perifraze 39, 111, 149  
 glagolski časi 21, 23, 37, 77, 123, 144,  
 149, 154  
 glagolski premik 29  
 glagolski vid 21, 111  
 glasovi 52, 106  
 gotščina 16  
 govor  
     berberski ~i 210  
     celjski ~ 106  
     trentinski ~ 208  
     triestinski ~ 20  
 gramatikalizacija 168, 218  
 grščina 5, 12, 16, 56, 64, 88, 190  
     moderna ~ 189  
 hetitščina 215  
 hrvaščina 7, 8, 48  
     moliška ~ 87  
 identiteta  
     evropske ~e 187  
     jezikovne ~e 187  
 idiomi 4, 84, 183, 193  
 imperfekt - gl. nedovršnik  
 implicitna zveza 167  
 intonacija 26, 153  
 islandščina 52, 77, 128  
 istroromunščina 46  
 italianizmi  
     ~ v hrvaščini 8  
     ~ v ladinščini 9  
     ~ v slovenščini 148  
 italijanščina 3, 21, 22, 36, 42, 43,  
 53, 54, 55, 68, 91, 103, 117,  
 118, 119, 123, 131, 144, 145,  
 146, 166, 167, 168, 169, 170,  
 171, 172, 173, 179, 202, 223, 225  
 izgovorjava 128, 138  
 izpeljanke 79, 197, 204, 207, 214  
 izposojenke 8, 9, 43, 82, 116, 148, 211  
 izražanje  
     govorno ~ 25, 26, 139, 146  
     pisno ~ 125, 126  
 izumiranje jezikov 120

périphrases verbales  
 temps verbaux  
  
 déplacement d'un verbe  
 aspect  
 sons  
 gotique  
 parler  
     ~s berbères  
     ~ de Celje  
     ~ de Trentino  
     ~ triestin  
 grammaticalisation  
 grec  
     ~ moderne  
 hitite  
 croate  
     ~ de Molise  
 identité  
     ~s européennes  
     ~s linguistiques  
 expressions idiomatiques  
 imparfait  
 connexion implicite  
 intonation  
 islandais  
 istroroumain  
 italianismes  
     ~ en croate  
     ~ en ladin  
     ~ en slovène  
 italien  
  
 prononciation  
 mots dérivés  
 emprunts  
 expression  
     ~ parlée  
     ~ écrite  
 disparition des langues

japonsčina 107, 158, 211, 222  
 Jespersenov krog 219  
 jeziki 164
 

- anatolski ~ 159
- baltski ~ 159
- indoevropski ~ 10, 76, 96, 159, 195
- romanski ~ 65, 114, 116, 208
- slovanski ~ 17, 159

 jeziki v stiku 8, 9, 43, 82, 116, 147, 148, 177, 178, 179, 180, 210, 211  
 jezikoslovci
 

- hrvaški ~ 113
- slovenski ~ 71, 73, 135, 137

 jezikovna biologija 62  
 jezikovna družina 196  
 jezikovna norma 25, 189  
 jezikovna raznolikost 49
 

- znotrajjezikovna ~ 161

 jezikovna vitalnost 162  
 jezikovne ravni 143  
 jezikovni atlas 9, 46, 95, 132, 180  
 kalki 178  
 karantanščina 82  
 kastiljščina 38, 185  
 kayardild (avstralski jezik) 209  
 kazala 227  
 keltščina 1, 104  
 klitike 84  
 književnost
 

- angleška ~ 191
- avstrijska ~ 186
- francoska ~ 141, 143
- hrvaška ~ 7
- italijanska ~ 42, 188
- ladinska ~ 67
- slovenska ~ 57, 182
- španska ~ 150, 151, 152, 186

 koherenca 3, 167  
 količinski določevalnik 27  
 kolokacije 105, 109, 129, 157  
 kontekst
 

- jezikovni ~ 169
- situacijski ~ 98

 japonais  
 cycle de Jespersen  
 langues
 

- ~ anatoliennes
- ~ baltiques
- ~ indo-européennes
- ~ romanes
- ~ slaves

 langues en contact  
 linguistes
 

- ~ croates
- ~ slovènes

 biologie du langage  
 famille de langues  
 norme (linguistique)  
 variation (linguistique)
 

- ~ intralinguistique

 vitalité des langues  
 niveaux de langue  
 atlas linguistique  
 calques  
 vieux slave alpin  
 castillan  
 kayardild (langue australienne)  
 index  
 celtique  
 enclitiques  
 littérature
 

- ~ anglaise
- ~ autrichienne
- ~ française
- ~ croate
- ~ italienne
- ~ ladine
- ~ slovène
- ~ espagnole

 cohérence  
 quantificateur  
 collocations  
 contexte
 

- ~ linguistique
- ~ situationnel

- korziščina 41  
 krajevna imena 14, 61, 82, 112, 175, 177  
 ladinščina 9, 67, 179  
 latinizmi v romanskih jezikih 116  
 latinščina 11, 18, 19, 37, 50, 64, 65,  
 66, 89, 90, 114, 122, 168  
 lažni prijatelji 108  
 leksikografija – gl. slovaroslovje  
 leksikologija – gl. besedoslovje  
 makedonščina 97  
 maroška arabščina 210  
 medkulturna pedagogika 187  
 medkulturno sporazumevanje 99, 100, 187  
 mestnik 76  
 metafora 78  
 morfemi  
 besedotvorni ~ 109, 157, 210, 213  
 prosti ~ 208  
 morfologija – gl. oblikoslovje  
 morfomi 209  
 morfosintaksa – gl. oblikoskladnja  
 naglaševanje 83, 207  
 nanašanje 98  
 naratologija 141, 143, 144, 149  
 narečje  
 galoitalska ~a 180  
 korziška ~a 92  
 ~ Arenzana 173  
 nemško ~ v Italiji (Sappada) 134  
 sicilijanska ~a 180  
 slovenska ~a 132, 147, 178  
 rezijsko ~ 177  
 tersko ~ 177  
 nedoločnik 56, 191  
 nedovršnik 21, 23, 37  
 nemščina 30, 35, 100, 129, 155, 156,  
 157, 179, 186, 192, 205, 213, 221  
 avstrijska ~ 186
- corse  
 toponymie  
 ladin  
 latinismes dans les langues romanes  
 latin  
 faux amis  
 lexicographie  
 lexicologie  
 macédonien  
 arabe marocain  
 pédagogie interculturelle  
 communication interculturelle  
 locatif  
 métaphore  
 morphèmes  
 ~ dérivationnels  
 ~ libres  
 morphologie  
 morphomes  
 morphosyntaxe  
 accentuation  
 référence  
 narratologie  
 dialecte  
 ~s gallo-italiques  
 ~s de Corse  
 ~ d'Arenzano  
 ~ allemand en Italie (Sappada)  
 ~s siciliens  
 ~s slovènes  
 ~ resian  
 ~ du Val de la Tour  
 infinitif  
 imparfait  
 allemand  
 ~ autrichien

neologizmi – gl. novotvorjenke	néologismes
novotvorjenke 198	néologismes
oblikoskladnja 147	morphosyntaxe
naravna ~ 2, 28, 44, 102, 212	~ naturelle
oblikoslovje 159, 168, 200, 206, 209, 210,	morphologie
211, 216, 218, 226	~ marginale
marginalno ~ 201, 202, 203, 205	~ naturelle
naravno ~ 74	~ prototypique
prototipsko ~ 203	morphologie canonique
oblikoslovna kanoničnost 206	modèle morphologique »parsli«
oblikoslovni model »parsli« 206	anthroponimie
osebna imena 176	marqueurs
označevalci	~ discursifs
diskurzni ~ 146	~ de l'infinitif
nedoločniški ~ 191	~ de direction
~ smeri 208	~ de négation
~ zanikanja 219	parfait
perfekt - gl. dovršnik	discours poétique
pesniški diskurz 150	pléonasme
pleonazem 88	sous-titrage
podnaslavljanje 192	constructions réductives
podvojitvene zgradbe 203	
pogled	anticipation
~ naprej 91, 144	rétropection
~ nazaj 144	subordonnée de condition
pogojni odvisnik 89	e muet
polglasnik 79	discours indirect libre
polpremi govor 42, 144	construction détachée
polstavek	~ participiale
deležniški ~ 45	~ infinitive
nedoločniški ~ 56, 190, 191, 192	sémantique
pomenoslovje 48, 177, 193	~ textuelle
besedilno ~ 155	terminologie maritime
pomorska terminologija 8	discours rapporté
poročani govor 6	phrase
poved	~ simple
enostavčna ~ 53	~ impérative
velelna ~ 145	

- povezovalci 24  
 praslovanščina 215  
 predlog 105, 109  
 prehodnost 165  
 prevajanje 55, 100, 108, 129, 192  
     literarno ~ 3, 85, 105, 182, 186  
 prevodne ustreznice 59  
 prevodoslovje 99  
 pridevnik 215, 216  
 primerjalno jezikoslovje 15, 16, 17,  
     26, 29, 31, 34, 38, 39, 45, 51, 57,  
     80, 81, 86, 96, 104, 107, 109, 127,  
     139, 154, 157, 159, 165, 167, 195,  
     196, 204, 205, 215, 218, 220, 226  
 primerniške zgradbe  
     predložne ~ 194  
     vezniške ~ 194  
 prislov 92, 158  
 prozodija izrekanja 139  
 recenzije 12, 13, 14, 31, 32, 33, 34,  
     35, 46, 47, 48, 49, 50, 61, 62, 63,  
     64, 65, 66, 67, 68, 69, 70, 115,  
     116, 117, 118, 119, 131, 132, 133,  
     134, 221, 222, 223, 224, 225, 226  
 retoromanščina 179  
 rodilnik 44  
 romunščina 17  
 rusizmi v italijanščini 43  
 ruščina 108  
 samijski jezik 78  
 samostalnik 215, 216  
 samostalniške manjšalnice 204  
 semantika – gl. pomenoslovje  
 sintagma 93  
     generična imenska ~ 97  
 sintaksa – gl. skladnja  
 skladnja 18, 29, 45, 48, 50, 53, 64, 66, 80,  
     89, 119, 122, 194, 216, 217, 219, 226  
 naravna ~ 56, 121, 138, 165, 190, 212  
 sklanjatev 44, 76, 159  
 slam 205  
 slovarji 116
- connecteurs  
 proto-slave (slave commun)  
 préposition  
 transitivité  
 traduction  
     ~ littéraire  
 équivalents  
 traductologie  
 adjectif  
 linguistique comparée
- constructions comparatives  
     ~ prépositionnelles  
     ~ conjonctionnelles  
 adverbe  
 prosodie de l'énonciation  
 recensions
- rétoroman  
 génitif  
 roumain  
 russismes en italien  
 russe  
 same  
 nom  
 noms diminutifs  
 sémantique  
 syntagme  
     ~ nominal générique  
 syntaxe  
 syntaxe  
     ~ naturelle  
 déclinaison  
 slam  
 dictionnaires



dvojezični ~ 59, 131	~ bilingues
enojezični ~ 4, 32, 110, 130	~ monolingues
uporaba ~ev 60, 130	emploi des ~
slovaroslovje 4, 32, 59, 110, 116, 131, 171	lexicographie
slovenščina 3, 26, 51, 55, 57, 59, 80, 82, 83, 84, 85, 100, 105, 107, 108, 109, 112, 127, 129, 131, 139, 147, 148, 154, 157, 177, 178, 182, 193, 194, 198, 204, 211, 218, 219, 220	slovène
slovnica 63, 68, 130	grammaire
historična ~ 168, 184	~ historique
opisna ~ 13	~ descriptive
primerjalna ~ 34	~ comparée
slovnične kode 110	codes grammaire
sociolingvistika 120, 171	sociolinguistique
splošno jezikoslovje 75	linguistique générale
spojenke 201, 202	amalgames lexicaux (mots-valises)
spol 171, 211	genre
ženski ~ 195	~ féminin
spontani govor 26, 139	discours spontané
stara cerkvena slovanščina 81	vieux slave
stara indijščina 102, 212	vieil indien
stavek	proposition
nikalni ~ 219	~ négative
vprašalni ~ 29, 121	~ interrogative
stereotipi 171	stéréotypes
stilistika 16, 154	stylistique
strokovni jezik	langue (langage) de spécialité
jezik umetnosti 103	langage de l'art
politični ~ 54, 163	~ politique
pravniški ~ 22, 129	~ juridique
reklamni ~ 35, 156	~ publicitaire
španščina	espagnol
13, 38, 111, 150, 183, 184, 185, 186	~ d'Amérique latine
ameriška ~ 149	
število 211	nombre
»tabula gratulatoria« 72, 136	»tabula gratulatoria«
teorija jezika 161	théorie du langage

teorija jezikovne naravnosti  
2, 10, 28, 44, 56, 74, 88, 100, 101,  
102, 121, 138, 165, 190, 212  
večjezičnost 187  
velelnik 217  
vezanje 138, 216  
vezljivost 109, 157  
vzročniške zgradbe 107  
zaimek 44, 92  
    kazalni ~ »ipse« 168  
    nikalni ~ 80  
    osebni ~ 19  
    oziralni ~ 40  
zanikanje 80, 219  
zemljepisno jezikoslovje 9, 62, 132, 180  
zlogovanje 79  
zloženske 74, 199, 225  
zvočniško izglasje 106

théorie du langage naturel

plurilinguisme  
impératif  
liaison  
valence  
constructions causatives  
pronom  
    ~ démonstratif »ipse«  
    ~ négatif  
    ~ personnel  
    ~ relatif  
négation  
géographie linguistique  
syllabation  
mots composés  
sonantes en final de mot

## VSEBINA – SOMMAIRE

### PRÉSENTATION

Predstavitev .....3

Michel Roché

PRESSION LEXICALE ET CONTRAINTES PHONOLOGIQUES DANS LA DÉRIVATION  
EN *-AIE* DU FRANÇAIS

Leksikalni vpliv in fonološke omejitve v izpeljavi s pripono *-aie* v francoščini .....5

Irena Stramljič Breznik/Ines Voršič

WORD-FORMATIONAL PRODUCTIVITY OF THE SLOVENE LANGUAGE IN THE CASE  
OF SPORTS NEOLOGISMS

Besedotvorna produktivnost slovenščine na primeru novotvorjenk  
s področja športa .....23

Alexandra Bagasheva

BEYOND DICHOTOMIES: ON THE NATURE AND CLASSIFICATION OF COMPOUND  
VERBS IN ENGLISH

Onstran dihotomij: o naravi in tipologiji glagolskih zloženek v angleščini .....39

Maris Camilleri

ISLAND MORPHOLOGY: MORPHOLOGY'S INTERACTIONS IN THE STUDY  
OF STEM PATTERNS

Otoška morfologija: vloga morfoloških interakcij pri raziskovanju  
vzorcev osnove .....65

Arnaud Léturgie

UN CAS D'EXTRAGRAMMATICALITÉ PARTICULIER : LES AMALGAMES  
LEXICAUX FANTAISISTES

Poseben primer izvensistemskosti: izmišljeni leksikalni amalgami .....87

Silvia Cacchiani

ON UNFAMILIAR ITALIAN LEXICAL BLENDS FROM NAMES AND NOUNS

O manj pogosti vrsti leksikalnih spojenk iz imen in samostalnikov .....105

Thomas Schwaiger

THE RELATION BETWEEN PROTOTYPICAL AND MARGINAL MORPHOLOGY:  
THE CASE OF REDUPLICATIVE CONSTRUCTIONS

Razmerje med prototipsko in marginalno morfologijo: primer  
reduplikativnih konstrukcij .....121

Eva Sicherl/Andreja Žele NOMINAL DIMINUTIVES IN SLOVENE AND ENGLISH Samostalniške manjšalnice v slovenščini in angleščini .....	135
Marie-Anne Berron /Marie Mouton LA MORPHOLOGIE MARGINALE DANS LE SLAM: COMPARAISON INTERLINGUALE FRANCE vs ALLEMAGNE Marginalna morfologija v slamu: francosko-nemška medjezikovna primerjava .....	143
Géraldine Walther MEASURING MORPHOLOGICAL CANONICITY Merjenje morfološke kanoničnosti .....	157
Tatjana Marvin PHASES AT THE WORD LEVEL Faze na besednem nivoju .....	181
Patrizia Cordin FROM VERBAL PREFIXES TO DIRECTION/RESULT MARKERS IN ROMANCE Od glagolskih predpon do označevalcev smeri/rezultata v romanskih jezikih .....	201
Erich R. Round MORPHOMES AS A LEVEL OF REPRESENTATION CAPTURE UNITY OF EXPONENCE ACROSS THE INFLECTION-DERIVATION DIVIDE Morfomi kot predstavitvena ravnina zajamejo enotnost izraznih podob ne glede na razkol med oblikoslovjem in besedotvorjem .....	217
Georgia Zellou MOROCCAN ARABIC BORROWED CIRCUMFIX FROM BERBER: INVESTIGATING MORPHOLOGICAL CATEGORIES IN A LANGUAGE CONTACT SITUATION Berberski cirkumfiks v maroški arabščini: raziskovanje morfoloških kategorij na primeru jezikov v stiku .....	231
Chikako Shigemori Bučar CREATIVE COMPETENCE IN BORROWINGS: WORDS OF JAPANESE ORIGIN IN SLOVENE Ustvarjalna kompetenca v izposojanju: besede japonskega izvora v slovenščini .....	245
Varja Cvetko Orešnik NATURALNESS: SOME SYNCHRONIC DATA OF OLD INDIAN Naravno jezikoslovje: nekaj sinhronih podatkov iz stare indijščine .....	263

Douglas Lightfoot	
TESTING THE SUFFIXOIDIZATION OF GERMAN <i>-MANN</i> ‘-MAN’	
Preverjanje sufiksoidizacije nemškega »-mann« ‘moški’ .....	273
Javier E. Díaz Vera	
DERIVATION IN A WORD-BASED MORPHOLOGY: ON THE ORIGIN OF OLD	
ENGLISH VERBS OF PERCEPTION, COGNITION AND EMOTION	
Izpeljevanje v besednoutemeljeni morfologiji: o izvoru	
staroangleških glagolov zaznave, spoznavanja in čustev .....	285
Metka Furlan	
THE ARCHAIC WORD-FORMATIONAL RELATIONSHIP BETWEEN THE	
PROTO-SLAVIC NOUN * <i>POL’E</i> AND THE HITTITE ADJECTIVE <i>PALHI-</i>	
Arhaično besedotvorno razmerje med praslovanskim samostalnikom	
* <i>pol’e</i> in hetitskim pridevnikom <i>palhi-</i> .....	291
Marc Plénat/Camille Plénat	
LA LIAISON DE L’ADJECTIF SUR LE NOM EN FRANÇAIS : MORPHOLOGIE,	
SYNTAXE, PHONOLOGIE	
Vežanje pridevnika na samostalnik v francoščini: morfologija,	
sintaksa, fonologija .....	299
Janez Orešnik	
THE IMPERATIVE IN NATURAL SYNTAX	
Velelnik v naravni skladnji .....	317
Mojca Schlamberger Brezar	
LE GÉRONDIF ET LE PARTICIPE PRÉSENT ET LEUR ÉVOLUTION VERS	
LA GRAMMATICALISATION : ÉTUDE CONTRASTIVE DU SLOVÈNE ET DU FRANÇAIS	
Deležja in deležniki sedanjega časa in njihov razvoj v smeri	
gramatikalizacije: slovensko-francoska kontrastivna študija .....	333
Gašper Ilc	
JESPERSEN’S CYCLE IN SLOVENIAN	
Jespersenov krog v slovenščini .....	349
Mojca Smolej	
QU’EST-CE QUE LE SLOVÈNE ET LE FRANÇAIS ONT EN COMMUN ?	
LES ARTICLES DÉFINI ET INDÉFINI	
Kaj imata slovenščina in francoščina skupnega?	
Določni in nedoločni člen .....	365

COMPTE RENDUS, RÉCENSIONS, NOTES – POROČILA, OCENE, ZAPISI

Barbara Sandig. <i>Textstilistik des Deutschen</i> (Stojan Bračič) .....	377
Toshiko Yamaguchi. <i>Japanese Linguistics: An Introduction</i> (Magnús Pétursson) ..	383
Paul Videsott. <i>Padania scrittologica. Analisi scrittologiche e scrittometriche di testi in italiano settentrionale antico dalle origini al 1525</i> (Ludwig Fesenmeier) ....	385
Raffaella Bombi e Fabiana Fusco (a cura di) ... <i>SAND CARRIED BY A STREAM... Scritti in onore di Vincenzo Orioles</i> (Paolo di Giovine) .....	390
Jan Radimský. <i>Les composés italiens actuels</i> (Gregor Perko) .....	395
Dany Amiot, Walter de Mulder, Estelle Moline et Dejan Stosic (éds.). <i>Ars Grammatica. Hommages à Nelly Flaux</i> (Gregor Perko) .....	399
Metka Šorli Linguistica XLI-LI .....	403

LINGUISTICA LI

Založila  
Znanstvena založba Filozofske fakultete Univerze v Ljubljani  
Izdal  
Oddelek za romanske jezike in književnosti

Revue éditée par les  
Presses scientifiques de la Faculté des Lettres  
et publiée par le  
Département des Langues et Littératures Romanes

Za založbo – Responsable  
Andrej Černe  
Dekan Filozofske fakultete – Doyen de la Faculté des Lettres

Glavna in odgovorna urednica – Rédactrice en chef  
Martina Ožbot

Številko LI uredil – Numéro LI dirigé par  
Gregor Perko

Tajnica redakcije – Secrétaire de la rédaction  
Metka Šorli

Nasloviti vse dopise na:  
Prière d'adresser toute correspondance à :  
Martina Ožbot  
Filozofska fakulteta  
Oddelek za romanske jezike in književnosti  
Aškerčeva 2  
1000 Ljubljana  
Slovénie

[linguistica@ff.uni-lj.si](mailto:linguistica@ff.uni-lj.si)

Tel. : +386 1 241 13 98  
Fax : +386 1 425 93 37

Naklada: 450 izvodov – Tirage : 450 exemplaires

Računalniški prelom – Mise en page  
Biro Stara Ljubljana, d. o. o.  
Rožna ulica 9, 1000 Ljubljana

Tisk – Impression  
Birografika BORI, d. o. o.  
Linhartova cesta 1, 1000 Ljubljana

Cena: 17 € – Prix : 17 €

