

# LINGUISTICA LVI

SODOBNE SMERNICE V  
TVORBENEM JEZIKOSLOVJU

Current Trends in Generative Linguistics

# LINGUISTICA

## LVI

**SODOBNE SMERNICE V  
TVORBENEM JEZIKOSLOVJU**

**Current Trends in Generative Linguistics**

Ljubljana 2016

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 LVI uredili – Responsable du numéro LVI  
Gašper Ilc, Frančiška Lipovšek, Tatjana Marvin, Andrej Stopar

Uredniški svet  
Janez Orešnik, Stojan Bračič, Gregor Perko

Svetovalni odbor – Comité consultatif  
Gábor Alberti, Boban Arsenijević, Aleš Bičan, Olga Borik, Daniel Buring,  
Guglielmo Cinque, Luka Crnič, Eva Csipak, Mojmír Dočekal, Jakub Dotlačil,  
Anna Gazdik, Hana Gruet-Skrabalova, Sabina Halupka-Rešetar, Veronika Hegedűs,  
Negin Ilkhanipour, Vadim Kimmelman, Judit Kleiber, Martin Krämer, Franc Marušič,  
Nataša Miličević, Petra Mišmaš, Nicola Munaro, Matic Pavlič, Ljiljana Progovac,  
Péter Rebrus, Branimir Stanković, Penka Stateva, Adrian Stegovec, Artur Stepanov,  
Balazs Suranyi, Ádám Szalontai, Aida Talić, Enikő Tóth, Barbara Vogt,  
Mary Ann Walter, Jacek Witkoś, Sarah Zobel, Rok Žaucer, Sašo Živanović

Izid revije je finančno podprla  
JAVNA AGENCIJA ZA RAZISKOVALNO DEJAVNOST RS

Sous les auspices de  
L'AGENCE NATIONALE SLOVÈNE POUR LA RECHERCHE SCINTIFIQUE

© Univerza v Ljubljani, Filozofska fakulteta, 2016. Vse pravice pridržane.

© Université de Ljubljana, Faculté des Lettres, 2016. Tous droits réservés.



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. / To delo je ponujeno pod licenco Creative Commons Priznanje avtorstva-Deljenje pod enakimi pogoji 4.0 Mednarodna licenca

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

81-116.6(082)

SODOBNE smernice v tvorbenem jezikoslovju = Current trends  
in generative linguistics / [uredili Gašper Ilc ... et al.]. - Ljubljana :  
Znanstvena založba Filozofske fakultete, 2016. - (Linguistica, ISSN 0024-3922 ; 56)

ISBN 978-961-237-885-1  
1. Vzp. stv. nasl. 2. Ilc, Gašper  
287997184

## SOMMAIRE – VSEBINA

PREFACE .....	7
 Olga Borik CONSTRAINTS ON THE POSITION AND INTERPRETATION OF BARE SINGULAR INDEFINITES IN RUSSIANN Omejitve pozicije ter interpretacije golih samostalnikov v edninski nedoločni obliki v ruščini .....	9
 Željko Bošković WHAT IS SENT TO SPELL-OUT IS PHASES, NOT PHASAL COMPLEMENTS Na točki loči se ločijo faze, ne fazna dopolnila .....	25
 Daniel Buring DISCONTINUOUS FOCI AND UNALTERNATIVE SEMANTICS Prekinjeni fokusi in nealternativna semantika .....	67
 Guglielmo Cinque A SOURCE OF PARAMETRIC VARIATION IN THE LEXICON Vir parametrične raznolikosti v leksikonu .....	83
 Mojmír Dočekal, Jakub Dotlačil EXPERIMENTAL EVIDENCE FOR NEG-RAISING IN SLAVIC Eksperimentalni dokazi za dviganje nikalnice v slovanskih jezikih .....	93
 Judit Farkas, Gábor Alberti THE RELATIONSHIP BETWEEN (IN)ALIENABLE POSSESSION AND THE (THREE POTENTIAL) FORMS OF POSSESSED NOUNS IN HUNGARIAN Razmerje med (ne)odtujljivo svojilnostjo in (trema potencialnimi) oblikami posedovanih samostalnikov v madžarščini .....	111
 Hana Gruet-Skrabalova YES OR NO, OR HOW TO ANSWER A NEGATIVE QUESTION Da ali ne, ali kako odgovoriti na nikalno vprašanje .....	127

Negin Ilkhanipour TENSE AND MODALITY IN THE NOMINAL DOMAIN Čas in modalnost v samostalniški domeni .....	143
Judit Kleiber, Gábor Alberti, Veronika Szabó THE INTENSIONAL PROFILES OF FIVE HUNGARIAN IMPERATIVE SENTENCE TYPES Intenzionalni profili petih madžarskih velelnih stavčnih tipov .....	161
Nedžad Leko, Nermina Čordalija, Ivana Jovović, Nevenka Marijanović, Lidija Perković, Dženana Telalagić, Amra Bešić, Midhat Šaljić STRATEGIES OF PARTICIPLE AGREEMENT WITH CONJOINED SUBJECTS IN BOSNIAN/CROATIAN/SERBIAN Strategije ujemanja deležnika s koordiniranimi osebki v bosanskem/hrvaškem/srbskem jeziku .....	173
Marijana Marelj IN THE SOUTH SLAVONIC GARDEN: LANDSCAPING THE LANDSCAPE OF ARGUMENTS AND NON-ARGUMENTS V južnoslovanskem vrtu: risanje pokrajine glagolskih argumentov in ne-argumentov .....	193
Nicola Munaro LOCALIZING CONDITIONAL CLAUSES IN THE LEFT PERIPHERY: EVIDENCE FROM MULTIPLE COMPLEMENTIZER CONSTRUCTIONS IN ROMANCE Skladenjski položaj in razvrstitev pogojnih odvisnikov v levi periferiji z vidika romanskih večvezniških struktur .....	211
Matic Pavlič SIGN ORDER IN SLOVENIAN SIGN LANGUAGE LOCATIVE CONSTRUCTIONS Znakovni red v krajevnih izrazih slovenskega znakovnega jezika .....	225
Péter Rebrus, Miklós Törkenczy TYPES AND DEGREES OF VOWEL NEUTRALITY Vrste in stopnje samostalniške nevtralnosti .....	239
Artur Stepanov, Manca Mušič, Penka Stateva ASYMMETRIES IN SUB-EXTRACTION OUT OF NP IN SLOVENIAN: A MAGNITUDE ESTIMATION STUDY Nesomernosti pri premiku iz slovenske samostalniške zveze: študija po metodi ocene magnitude .....	253

Ildikó Emese Szabó LOW VOWEL “LENGTHENING” IN HUNGARIAN »Daljšanje« nizkih samoglasnikov v madžarščini .....	273
Enikő Tóth, Péter Csátár INDEXICAL DEMONSTRATIVES AND IDENTIFICATIONAL FOCUS IN HUNGARIAN Indeksialni kazalni zaimki in identifikacijski fokus v madžarščini .....	281
Lucas Tual AMBIGUOUS ADJECTIVES IN FRENCH: THE CASE OF “GROS” WHEN COMBINED WITH DEVERBAL NOUNS Dvoumni pridevniki v francoščini: primer pridevnika „velik“ v kombinaciji z izglagolskimi samostalniki .....	293
Mary Ann Walter LEXICAL STRATA AND VOWEL (DIS)HARMONY: THE TURKISH TRANSFORMATION OF A BALKAN HYPOCORISTIC Leksikalni sloji in samoglasniška (dis)harmonija: sprememba balkanske hipokoristike v turščini .....	309
Jacek Witkoś, Dominika Dziubała-Szrejbrowska SOME ASPECTS OF AGREEMENT WITH NUMERAL PHRASES IN POLISH Nekateri vidiki ujemanja s števnimi zvezami v poljščini .....	321
Sarah Zobel, Eva Csipak CONDITIONAL ANTECEDENTS CONTAINING THE GERMAN DISCOURSE PARTICLE <i>DENN</i> : A CORPUS STUDY Nemški pogojni odvisniki z diskurzivnim členkom <i>denn</i> : korpusna analiza .....	345
Sarah Zobel ON THE (IN)DEFINITENESS OF IMPERSONAL PRONOUNS O (ne)določnosti neosebnihi zaimkov .....	363



## PREFACE

This volume of *Linguistica* is dedicated to the research in the field of generative linguistics. It contains a careful selection of academic articles which focus on contemporary developments in the generative approach to linguistic description. Parts of these articles were also presented at the 2015 *Syntax, Phonology and Language Analysis Conference* (SinFonIJA 8), held at the University of Ljubljana Faculty of Arts from September 24<sup>th</sup>–26<sup>th</sup> 2015, and at the 2014 *Syntax, Phonology and Language Analysis Conference* (SinFonIJA 7), held at the University of Graz from September 25<sup>th</sup>–27<sup>th</sup> 2014.

The 22 papers are organized alphabetically. They are from various subfields of (generative) linguistics: phonology (Rebrus and Törkenczy; Szabó; Walter), semantics (Büring; Kleiber; Alberti and Szabó; Zobel), syntax (Bošković; Munaro; Witkoś and Dziubała-Szrejnbrowska), syntax-semantics interface (Borik; Cinque; Gruet-Skrabalova; Ilkhanipour; Marelj), syntax-morphology interface (Farkas and Alberti; Tual) and experimental linguistics (Tóth and Csatár (experimental pragmatics), Csipak and Zobel (experimental semantics, corpus study), Pavlič (sign language syntax), Dočekal and Dotlačil; Leko et al.; and Stepanov, Mušič and Stateva (experimental syntax)).

The authors come from universities and research institutes in 14 different countries: Austria (1 author), Bosnia and Herzegovina (8 authors), Czech Republic (1 author), France (2 authors), Germany (2 authors), Hungary (9 authors), Italy (2 authors), Iran (1 author), the Netherlands (2 authors), Poland (2 authors), Slovenia (4 authors), Spain (1 author), Turkey (1 author), and the United States of America (2 authors).

Every paper received sets of comments from two or three reviewers and was further read by the editors. We would like to seize the opportunity and thank the following list of reviewers: Gábor Alberti, Boban Arsenijević, Aleš Bičan, Olga Borik, Stojan Bračić, Daniel Büring, Guglielmo Cinque, Luka Crnič, Eva Csipak, Mojmír Dočekal, Jakub Dotlačil, Anna Gazdik, Hana Gruet-Skrabalova, Sabina Halupka-Rešetar, Veronika Hegedűs, Negin Ilkhanipour, Vadim Kimmelman, Judit Kleiber, Martin Krämer, Franc Marušič, Nataša Miličević, Petra Mišmaš, Nicola Munaro, Matic Pavlič, Ljiljana Progovac, Péter Rebrus, Branimir Stanković, Penka Stateva, Adrian Stegovec, Artur Stepanov, Balazs Suranyi, Ádám Szalontai, Aida Talić, Enikő Tóth, Barbara Vogt, Mary Ann Walter, Jacek Witkoś, Sarah Zobel, Rok Žaucer, and Sašo Živanović.

Ljubljana, November 2016

The Editors







## CONSTRAINTS ON THE POSITION AND INTERPRETATION OF BARE SINGULAR INDEFINITES IN RUSSIAN\*\*

### 1. INTRODUCTION

It is quite uncontroversial that in languages which do not have overt articles, nominal phrases that appear bare can, in principle, be interpreted in various ways, including as definite, as in (1a) and (1b), as indefinite, as in (1b), or as generic/kind, as in (1c) below:

- (1) a) *Poezd prišel.*  
       train.nom arrived  
       ‘The/#A train arrived.’
- b) *Prišel poezd.*  
       arrived train.nom  
       ‘The/A train arrived.’
- c) *Poezd kak sredstvo peredviženija očen’ udoben.*  
       train.nom as means transportation.gen very convenient  
       ‘The train as a means of transport is very convenient.’

For Russian, this view has always been supported by, for instance, traditional grammars. However, in recent semantic literature, an influential proposal has been made, according to which bare nominals in Russian or Hindi – both languages with no overt articles – can only be interpreted as kind or definite (Dayal 2004), whereas bare singular indefinites in these languages virtually do not exist (at least not in the subject position) or are very restricted. In this paper, I will demonstrate that bare singular indefinites appear quite regularly in Russian, contra Dayal’s (2004) proposal. However, the distribution of these nominals is, indeed, subject to certain restrictions, and I will examine in detail the nature of this restriction in the preverbal subject position. One of the previous analyses of this restriction proposed by Geist (2010) relies on the claim that bare singular indefinites are never specifically interpreted. I will argue that this analysis must be modified, since a specific interpretation does, in principle, quite regularly appear with bare singulars in the object position.

---

\* olga.borik@gmail.com

\*\* This research has been funded by grants FFI2014-52015-P (awarded by the Spanish Ministerio de Economía y Competitividad) and 2014 SGR 1013 (awarded by the Generalitat de Catalunya). I am very grateful to two anonymous reviewers for their insightful comments. All remaining errors are mine.

In my paper, I will focus only on bare singular nominals. There is much cross-linguistic evidence that supports the view that bare singulars could (and/or should) be treated differently, and separately, from bare plurals. For instance, it is well-known that in English, a language with overt articles, bare plurals can have a generic reading (cf. Carlson 1980). This is also a reading that arises freely and prominently with bare plurals in Russian. Bare singulars, on the other hand, appear to have a relatively restricted distribution in languages with overt articles (cf., for instance, Stvan 1998 for English; Munn & Schmitt 2005 for Brazilian Portuguese; Doron 2003 for Hebrew; Borthen 2003 for Norwegian; Espinal & McNally 2011 for Spanish and Catalan; de Swart 2015 for Dutch, English and French) and often receive an incorporation analysis which presupposes that they are not fully referential. In languages without articles, however, bare singulars (henceforth BSGs) possess quite different properties, and since their referential uses are often unmarked, they naturally give rise to a wide range of interpretations in many contexts, as was illustrated in (1) above.

The paper is organised as follows. Section 2 introduces empirical data and gives a brief overview of indefinite uses of BSGs in Russian. Section 3 is devoted to the main theoretical notions that are relevant for the analysis of the facts presented in section 2, *i.e.* topicality and specificity. In section 4, I present a proposal for BSGs in subject/topic position and discuss several consequences of this proposal. Section 5 concludes the paper.

## 2. INDEFINITE BARE SINGULARS IN RUSSIAN

Let me begin this section by illustrating that Russian BSGs can, indeed, have an indefinite reading, contrary to Dayal's (2004) proposal for article-less languages, where indefinite (existential) readings are supposed to be derived from a kind reading. Consider the following Russian example.

- (2) *V kazhdom dome igral rebenok.*  
 in every house played child.NOM  
 'A child (a different one) was playing in every house.'

In a distributive context, like the one given in (2), a nominal phrase cannot be interpreted as either a kind reading, since 'singular' kinds generally cannot facilitate access to individuals (cf. Krifka et al. 1995) and hence cannot distribute, or as a definite reading, since this would lead to an implausible interpretation that there is a unique child playing in every house. Hence, it must be indefinite. Dayal's crucial examples for the absence of an indefinite interpretation with bare singulars in Russian are also based on distributive contexts,<sup>1</sup> but, as (2) illustrates, distributive sentences do not rule out singular indefinites on a regular basis.

1 Dayal's (2004) examples include BSGs in the preverbal position. Later in the paper, I will argue that BSGs are not readily used as preverbal subjects for other reasons. For some other observations and considerations concerning Dayal's specific examples see, for instance, Bronnikov (2004).

Another argument in favor of the view that Russian bare singulars do have a proper indefinite interpretation comparable to English *a*-indefinites (i.e., singular nominals with an indefinite determiner) is based on the scope ambiguities that Russian bare singulars regularly exhibit in the object position.<sup>2</sup> This ambiguity is illustrated in the following example:

- (3) *Vasja hočet ženit'sja na kinozvezde,*  
 Vasja wants marry on movie-star  
 'Vasja wants to marry a movie star...'  
 a) *no ne možet najti podhodjasčuju.* want > Indef  
 but not can find suitable  
 ...but cannot find a suitable one.'  
 b) *no my eje esče ne videli.* Indef > want  
 but we her yet not saw  
 ... but we haven't met her yet.'

As shown in (3), a BSG indefinite in an opaque context (i.e., in a complement clause of a propositional attitude verb) can have two readings, which correspond to two different interpretations of the sentence. One reading, illustrated in (3a), is called *opaque*, in which case the indefinite is interpreted within the scope of the main verb. A *transparent* reading in (3b) is one in which the indefinite is outside the scope of the main verb. In this latter case, the indefinite is interpreted as referential, or specific, while in (3a) it is a purely quantificational indefinite. The ambiguity illustrated here is a classical test for specific and non-specific (existential) readings of indefinite noun phrases in English and, as (3) illustrates, Russian BSG indefinites are not really different from their English counterparts in allowing both readings.

Finally, let us look at yet another canonical indefinite environment, i.e., existential *there* sentences. The Russian counterpart of *there* sentences in English are formed by fronting a locative phrase, as in (4):

- (4) *V komnate ležal kover.*  
 in room lied carpet  
 'There was a carpet in the room.'

The example above has a BSG nominal phrase *kover* 'carpet', which can only be interpreted as an indefinite in the given context.

So far, I have demonstrated that BSGs in Russian freely appear in typical indefinite environments and, moreover, exhibit ambiguities just as bona fide indefinites in English. We can therefore conclude that BSGs in Russian do function as proper full-fledged indefinites.

---

2 Note that Dayal (2004) does not discuss BSGs in the object position, as in Hindi they can be pseudo-incorporated. A pseudo-incorporation analysis for object BSGs in Russian is not really a plausible option since regular accusative objects do not exhibit any properties of pseudo-incorporated nominals.

There are, however, certain distributional restrictions associated with indefinite BSgs in Russian. These restrictions primarily concern the preverbal subject position. In particular, indefinite subjects in Russian are usually postverbal, whereas BSgs in the preverbal subject position have a tendency to be interpreted as definite. The following example provides an illustration for the above restriction.

- (5) *V komnate bylo neskol'ko malen'kih detej, mal'čikov i devoček.*  
 in room were several small children boys and girls  
 'There were several small children in the room.'
- a) *#Devočka podošla ko mne i sprošila...*  
 girl.NOM came.up to me and asked...
- b) *Ko mne podošla devočka i sprošila...*  
 to me came.up girl.NOM and asked...  
 'A girl came up to me and asked...'

In this example, the first sentence in (5) sets a context, which makes sure that there are several boys and girls in the group of children. In this particular context, the continuation in (5a) with an intended indefinite interpretation of *devočka* 'girl' sounds odd, as the nominal subject has a clear preference for a definite interpretation: the sentence would be perfectly acceptable if there was one girl in the group. If, however, the BSg *devočka* 'girl' is used postverbally, as in (5b), an indefinite interpretation is entirely adequate.

These restrictions have not gone unnoticed in the literature. In particular, Geist (2010) proposes an analysis which relies on the unavailability of a specific interpretation for indefinite BSg nominals in Russian. The following points comprise the main components of her proposal. First, she assumes that all preverbal subjects in Russian are topics.<sup>3</sup> With respect to topics, she adopts Reinhart's (1981) proposal, according to which only strongly referential (*i.e.*, specific) indefinites can function as topics. Geist then argues that Russian BSgs cannot have a specific interpretation and that they are therefore excluded from the topic position. This is an explanation offered for the 'oddness' of (5a).

An immediate objection to this proposal is that, in general, BSg indefinites in Russian cannot solely be associated with a non-specific interpretation. This becomes evident if one takes into account examples such as (3) above and the ambiguity associated with BSgs in the object position.<sup>4</sup> This ambiguity can also be illustrated in a different context, for instance, for an indefinite under negation. Consider the following example:

- (6) *Vasja byl zloj potomu čto segodja on ne sdal ekzamen.*  
 Vasja was angry because that today he not pass exam  
 'Vasja was angry because today he didn't pass an exam.'

3 A more detailed explanation of the notion of topic is given in the next section.

4 Geist (2010) claims that this ambiguity does not exist. Native speakers that I have consulted, however, support my own judgment and agree that both (3a) and (3b) are fine.

Just as in (3), a BSg *ekzamen* ‘exam’ in (6) can stay in the scope of the negative operator, yielding an interpretation ‘it is not the case that Vasja passed an exam today’, or scope out of the negative operator and be interpreted specifically, *i.e.*, there was an exam that Vasja didn’t pass. Crucially, the data in (3) and (6) indicate that the reason why indefinite BSgs do not easily appear in a preverbal position cannot be associated with a ban on a specific interpretation for indefinite BSgs. Hence, it becomes necessary to seek an alternative explanation. Before offering such an explanation, I will clarify the basic theoretical notions on which I will build my case, namely, topicality and specificity.

### 3. THEORETICAL ASSUMPTIONS: THE NOTIONS OF TOPICALITY AND SPECIFICITY

#### 3.1. Topics

Although the concept of topicality has been around for many years and its relevance for linguistics is not a subject of any considerable debate, there are still many ways in which topics are understood. For concreteness, in this paper, the following view on topics is assumed. First of all, only *sentential* topics and not *discourse* topics are considered here. A sentential topic is what a sentence is *about*. This informal characterisation of topicality is known as the one in terms of *aboutness* and is adopted in Reinhart (1981) and Endriss (2009), among many others. The following example taken from Endriss (2009: 20) can provide an elementary illustration of a topic in a sense adopted here.

- (7) a) *Yesterday, Clarissa visited Dena.*  
b) *Yesterday, Dena was visited by Clarissa.*  
c) *The telephone rang.*

The sentence in (7a) is naturally understood as saying something about Clarissa, whereas (7b) is rather a statement about Dena. In this sense, Clarissa is a topic of (7a) and Dena is a topic of (7b).

From the point of view of information structure, a simple declarative sentence typically consists of two parts, a topic (an entity, in a broad sense) and a comment (saying something about this entity, elaborating on it). However, not all sentences have this structure; some sentences just present a piece of information as a single unit. This distinction is well-known and is usually stated in terms of categorical *vs.*thetic judgments, where only the former have topics, while the latter simply convey all the information in one step, without any division into a topic and a comment. An example of a thetic judgment is given in (7c) above. The behavior of indefinite BSgs in thetic judgements will be of special importance later in the paper.

The examples in (7a) and (7b) illustrate an intimate connection between the grammatical (syntactic) notion of subject and the information structural notion of topic, in the sense that very often subjects are also topics. This does not mean, however, that topic and subject are two equivalent notions: topic is used to describe the information

structure of a sentence, whereas subject is a purely grammatical category. In passive sentences or in left dislocation structures, topic and subject are clearly dissociated. However, the discussion in this paper is limited to topical subjects only, so for the purposes of this paper it is safe to assume that preverbal subjects in Russian categorical sentences with a neutral word order are by default topics in the sense described above.

Having briefly characterised the notion of topic, I will now turn to another controversial notion in the semantic/pragmatic literature, namely, specificity.

### 3.2 Specificity

In the most general terms, specificity is a notion that concerns different interpretations of indefinite noun phrases. The first discussions of the difference between specific and non-specific indefinites (cf. Karttunen 1968; 1969[1976]) were inspired by the observation that some indefinites can introduce discourse referents, while others do not have this capacity.

In later literature, specificity has been discussed in connection to referentiality (Fodor & Sag 1982), exceptional scope behavior (Reinhart 1997; Winter 1997), and presuppositionality (Geurts 2002[2010]). Which of these properties is essential to specificity is still a subject of debate, but what seems to emerge from the recent literature on specificity is that referential properties of indefinites are intimately connected to both their discourse status and their scopal properties (cf. von Heusinger 2011 for an overview).

The view on specificity assumed in this paper is based on the referential properties of indefinites. In particular, a specific indefinite has a particular referent, that is, the set of potential referents for an indefinite is limited to few or possibly one entity, whereas a non-specific indefinite does not refer to any entity. This type of ambiguity has already been illustrated above for Russian BSGs in (3) and (6), but let us now look at an example in English (from Dahl 1970):

- (8) *She wants to marry a man with a big bank account.*  
a) There is a man with a big bank account that she wants to marry.  
b) She wants there to be a man with a big bank account for her to marry.

The example in (8) is ambiguous between the senses in (8a) and (8b). (8a) conveys a specific interpretation of an indefinite, with an existential entailment, *i.e.*, there is a man that she wants to marry. (8b), on the other hand, does not entail that there is a man that she wants to marry and corresponds to a non-specific interpretation, when the indefinite does not refer to any particular man.

In Haspelmath's (1997) typology, specific indefinites can be of two types, known to the speaker and unknown to the speaker, as illustrated in (9) and (10):

- (9) Type 1. Specific, known to speaker:  
*Somebody called while you were away: guess who?*

(10) Type 2. Specific, unknown to the speaker:

*I heard **something**, but I couldn't tell what kind of sound it was.*

In both cases, an indefinite pronoun refers to a specific entity (a person or a thing), but in (9) the speaker can identify the referent, whereas in (10), the speaker is ignorant about the identity of the referent.

Finally, note that specific indefinites are different from definite nominal phrases in at least two respects. First, a specific indefinite does not trigger any uniqueness presupposition, *i.e.* the referent of an indefinite expression is not required to have a unique value shared by the speaker and the hearer. Second, indefinites in general are not anaphoric, *i.e.*, they are not required to refer to a previously mentioned or previously introduced discourse referent (cf. Heim 1982).

In Russian, specificity can be encoded overtly by means of so-called ‘indefinite pronouns’ (Academy Grammar of Russian 1982), which explicitly mark specificity distinctions:

(11) a) *Ja xoču koe-čto tebe skasat'.*

I want something you tell

‘I want to tell you something’/‘I have something to tell you.’

b) *Maša xočet vyjti замуž za kakogo-to/odnogo izvestnogo bankira.*

Maša wants marry prep some / one famous banker

‘Maša wants to marry a/some/one famous banker.’ (there is a specific banker)

c) *Maša xočet vyjti замуž za kakogo-nibud' izvestnogo bankira.*

Maša wants marry prep some famous banker

‘Maša wants to marry a/any famous banker.’ (there is no specific banker)

The interpretation of the pronoun *koe-čto* ‘something’ in the object position in (11a) is specific and known to the speaker: the speaker definitely knows what exactly s/he is about to say. In (11b), the specificity marker *kakogo-to* ‘some’ or *odnogo* ‘one’, used with a noun in the object position, indicates that there is a specific famous banker that the girl wants to marry, but in (11c) the interpretation of the whole nominal phrase with the marker *kakogo-nibud'* ‘some’ can only be non-specific.<sup>5</sup>

While indefinites with various specificity markers in Russian have received considerable attention in recent semantic literature (Bylinina & Testelec 2004; Yanovich 2005; Geist 2008; Ionin 2013; etc.), bare indefinites and their readings have not been investigated in detail. In particular, there is little research dedicated to the question of what kind of interpretation a bare indefinite can convey. I will again contrast the view advocated here with the analysis offered by Geist (2010). In this paper, I argue that BSG indefinites can indeed have a specific interpretation, whereas Geist’s (2010) proposal is based on the claim that they cannot. In section 2 (cf. the discussion of examples (3)

5 The type of specificity illustrated in (11) with overt markers is often called ‘epistemic’ specificity in relevant literature (cf. Ionin 2009). There is an intimate connection between scopal and epistemic specificity, which some accounts (for instance, Kratzer 1998) make explicit.



and (6)), I have demonstrated that, at least in object position, BSG indefinites exhibit regular scope ambiguities that are observed with bona fide indefinites in English.<sup>6</sup> This means that a non-specific interpretation is not the only one available for BSG nominals in Russian. In the following section, I present a proposal which takes this into account.

#### 4. A PROPOSAL AND ITS CONSEQUENCES

The proposal that I will put forward in this paper is that BSG indefinites in Russian are underspecified with respect to specificity. This characterisation is not reflected in a syntactic or semantic composition of BSG nominals themselves, although the ambiguity conveyed by a BSG is reflected in the logical form of a sentence. Specificity is a pragmatically oriented notion, so it is only reasonable to assume that it has no syntactic representation (unless overtly marked) and specific/non-specific phrases are only distinguished contextually.<sup>7</sup> Underspecification means that, in principle, BSGs can have both specific and non-specific interpretations, and that in many contexts where BSGs are found, they will be ambiguous between the two readings. This is exactly what happens in (3)/(6) from section 2 above, where a BSG in the object position can render various (non-)specific readings.

Now I will return to the restriction on BSG indefinites in the subject position.<sup>8</sup> Just like it is assumed in Geist (2010), I take preverbal subjects in Russian to be topics. Let me emphasise that this is not a ‘topic position’ in a syntactic sense, that is, I do not think that Russian preverbal subjects necessarily ‘raise’ to a specifier of a special topic

---

6 While it has been argued that there is no direct dependency between wide scope and specificity (cf. Enç 1991), the tendency of specific indefinites to take a wide scope remains, and the interpretation of the object in (3) does not leave any doubt that this nominal expression can have a specific referent.

7 There is an ongoing debate concerning the question of what would be an appropriate syntactic representation for nominal arguments in articleless languages. The proponents of the universal DP hypothesis (*e.g.*, Pereltsvaig 2006) argue for the presence of a null D in languages without articles, whereas those who advocate the parametrised DP hypothesis (*e.g.*, Bošković 2008) argue for a bare NP structure for nominal phrases in languages like Russian. There are, however, strong reasons to believe that there are nominal projections of different ‘sizes’ both in languages with and without articles and this syntactic difference is reflected in the interpretation of nominals. This view is advocated in Pereltsvaig (2006) and I support it, although I do not have the possibility of defending my position here. Based on this view, however, I assume that indefinite nominal arguments are represented as DPs with an underspecified D (cf. Ramchand & Svenonius 2008), so that both an indefinite and a definite interpretation can be derived with the same null element.

8 By now the reader might wonder if something is amiss in the argument since I shift freely from BSGs in object position to those that appear in subject position. The point I am making, however, remains valid: if BSGs cannot have a specific interpretation, it should be a property of a BSG itself, and not a property of the environment in which it appears. What I demonstrate is that the inavailability of a (non-)specific indefinite reading cannot be attributed to the structural properties of a nominal argument itself.

In relation to the subject/object asymmetry, a reviewer asked whether or not it is possible to topicalise an object argument in Russian. The answer is yes, and topicalised objects are actually interpreted definitely (cf. Erteschik-Shir 2013), as in (i):

projection. I simply assume that preverbal subjects are syntactically subjects (that is, they are found in a ‘standard’ subject position, like a SpecTP, for instance), but from the perspective of the information structure, they are topics, *i.e.* they name an entity about which the rest of the sentence says something. Sentential topics are not obligatory, whereas subjects (at least if we assume EPP) are structurally indispensable. Furthermore, I adopt the position defended in Reinhart (1981) that only specific indefinites can be topical.

The combination of these assumptions leads to the following picture with respect to BSgs in Russian. BSgs, as I have suggested above, are underspecified with respect to specificity. This means, under the definition of specificity adopted here, that these expressions do not necessarily have a referent, but that it can be established if additional information is provided. This information can be sentential or, possibly, contextual, but I will not consider contextual factors in a broad sense here. The point is that BSgs cannot function as sentential topics unless they are disambiguated and established as referring specifically.

Let us now go back to the context for which it was first shown that BSgs cannot freely occur as topics. The context was given in (5) in section 2 and is repeated below:

- (12) *V komnate bylo neskol’ko malen’kih detej, mal’čikov i devoček.*  
 in room were several small children boys and girls  
 ‘There were several small children in the room.’
- a) *#Devočka podošla ko mne i sprosila...*  
 girl came.up to me and asked...  
 b) *Ko mne podošla devočka i sprosila...*  
 to me came.up girl.nom. and asked...  
 ‘A girl came up to me and asked...’

Once again, an explanation that I propose for the oddness of (12a) is based on the claim that BSgs are underspecified with respect to specificity and the subject of (12a) should be interpreted as a topic, *i.e.* as having a specific referent. There are at least two

- 
- (i) (a professor, talking about his first lecture for a big audience)  
*Devušku ja zapomnil očen’ xorosho: ona zapisyvala každyj primer.*  
 girl.ACC I.NOM remembered very well: she copied every example  
 ‘The/\*A (certain) girl I remember very well: she copied every example.’

The sentence in (i) presupposes that there was only one girl in the whole audience. Thus, a specific indefinite interpretation is impossible for the topicalised bare object in (i); a specificity marker is obligatorily used to appropriately render such an interpretation, as in (ii):

- (ii) (a professor, talking about his first lecture for a big audience)  
*Odnu devušku ja zapomnil očen’ xorosho: ona zapisyvala každyj primer.*  
 one girl.ACC I.NOM remembered very well: she copied every example  
 ‘There was a girl that I remember very well: she copied every example.’

Thus, topicalised objects seem to be subject to the same type of restrictions as preverbal subjects (*cf.* the discussion of (5)).

possible ways to make the subject of (12a) specific: either by adding an overt specificity marker, or by enriching the descriptive content of the nominal phrase (cf. also Geist 2010 for this observation) to narrow down a set of possible referents. Both strategies lead to a full acceptance of a modified sentence in the same context, as demonstrated in the following two examples:

- (13) *Odna/kakaja-to devočka podošla ko mne i sprosila...*  
 one/some(spec.) girl came.up to me and asked...  
 ‘One/some girl came up to me and asked...’
- (14) *Devočka so slomanoj kukloj v ruke podošla ko mne i sprosila...*  
 girl with broken doll in hand came.up to me and asked...  
 ‘A girl with the broken doll in her hand came up to me and asked...’

To check some of the further consequences of the proposal, we will now examine the behaviour of indefinite BSgs in two other types of constructions, namely, as subjects of individual level (i-level) predicates and as subjects ofthetic jugements. The main reason for choosing these two constructions is that they have a (relatively) clear status with respect to the topic-comment structure. In particular, the subjects of i-level predicates have been argued to always be topics (cf. Chierchia 1995), whereasthetic jugements by definition lack topics (Ladusaw 1994). Therefore, there are two opposite expectations with respect to BSg indefinites in these constructions: they should be ruled out in combination with i-level predicates and available without any restrictions inthetic sentences.

First, I-level predicates should be considered. The following example from English can be used to illustrate that indefinite subjects of i-level predicates must be strongly referential:

- (15) *A fireman is altruistic.*

In this sentence, the indefinite subject can only have a generic or a specific (so-called ‘strong’ referential reading), but cannot refer to an arbitrary, non-specific fireman. If i-level predicates in Russian impose the same restrictions on the interpretation of their indefinite subjects (and there should be no a priori reason why they would not), one expects BSg indefinites to be ‘difficult’ in these contexts, as they are in the topic position of other categorical sentences. Once again, specificity markers should help to turn an underspecified BSg into a fully referential indefinite, which will be acceptable as a subject of an i-level predicate. Consider now the examples:

- (16) a) *Student byl smyšlenyj.*  
 student was smart  
 ‘The/#A student was smart.’
- b) *Odin student byl smyšlenyj.*  
 one student was smart  
 ‘A (specific) student was smart.’

As expected, a BSG in (16a) cannot really be interpreted as indefinite. The reason would presumably be the same one that rules out topic BSGs in episodic sentences: their underspecification with respect to specificity, and their inability to have a particular referent without any additional means. Once an overt specificity marker *odin* (one) is used in a modified version of the sentence, as in (16b), the indefinite nominal becomes acceptable.

Consider nowthetic judgments. By definition,thetic judgments are topicless, so one's expectation is that there is no restriction on the interpretation of indefinite nominal phrases inthetic judgments and, in particular, no restriction on BSGs in this type of sentences. These expectations are indeed supported by the data. If one considers a sentence with a BSG subject that is used in some typical context for athetic judgment, then, indeed, any restrictions on the interpretation of BSGs seem to disappear. This is illustrated in (17).

(17) Situation: All of a sudden, A and B see a big crowd gathered around an ambulance and a police car on a busy street.

A: *Čto slučilos', kak ty dumaeš?*  
 what happened how you think  
 'What do you think happened?'

B: *Mašina, naverno, kogo-nibud' sbila.*  
 car probably someone hit  
 'Probably a car hit somebody.'

In this example, a context is set that makes it highly implausible for a BSG in subject position in the B-answer to be interpreted as anything but an indefinite. Note that this indefinite can be both specific and non-specific, but in the absence of any additional (linguistic or extra-linguistic) information, the subject – *mašina*, or 'car' – tends to be interpreted as non-specific. A specific interpretation seems to be difficult to obtain without any specificity markers.<sup>9</sup>

To conclude this section, I will briefly address the question of motivating a connection between specificity and topics. The link between the two notions becomes more apparent once certain additional assumptions about topics are made. In particular, there is a view in relevant literature that considers a topic to be not a linguistic entity (*i.e.*, a phrase) in and of itself, but rather a denotation of this entity (Dahl 1974; Portner and Yabushita 1998; Endriss 2009). If this hypothesis is adopted, it naturally follows that only referential expressions can serve as topics, since non-referential indefinites do not denote entities. In other words, if a linguistic expression is of the type that does not or cannot have a particular referent, this expression cannot serve as a topic. If BSG indefinites are underspecified, the referent of a BSG indefinite in Russian cannot be unambiguously

9 The reasons why the pattern is such remain to be understood. One speculation is that a specific reading is difficult to obtain for a BSG precisely because specificity is one of the very few nominal categories that can be marked in Russian. A pragmatic principle could be responsible for a preference for specificity markers in those cases where they can be appropriately used.

established and consequently these expressions cannot be freely used in the topic position. In other words, the topic position does not tolerate underspecification.

## 5. CONCLUSION

In this paper, I have demonstrated that an indefinite interpretation is, indeed, available for BSgs in Russian (contra Dayal 2004) and that Russian indefinite BSgs are comparable to their canonical English counterparts with an indefinite article. In particular, BSgs can be interpreted both specifically and non-specifically, and can take a wide scope with respect to other scopal elements in a sentence. On the other hand, BSg indefinites do not freely appear in preverbal subject/topic position. I have argued that this is because a specific interpretation, or, in other words, a referential reading, required for a nominal in this position cannot be unambiguously established with BSgs, at least not without any additional sentential specification. Note that in the object position, where the information structure does not pose any special requirements, a BSg can remain underspecified, which results in two available interpretations of a sentence: one in which a BSg object is interpreted non-specifically, and another in which a BSg object has a specific reading.<sup>10</sup>

## References

- BOŠKOVIĆ, Željko (2008) “What Will You Have, DP or NP?” In: E. Elfner/M. Walkow (eds), *Proceedings of NELS 37*. Amherst, MA: GLSA, 101–114.
- BORTHEN, Kaia (2003) *Norwegian Bare Singulars*. PhD dissertation. Trondheim: NTNU.
- BRONNIKOV, George (2004) *A Critique of Dayal*. Ms. University of Texas at Austin.
- BYLININA, Lisa/Yakov G. TESTELEK (2004) “Sluicing-Based Indefinites in Russian.” In: S. Franks et al. (eds), *Formal Approaches to Slavic Linguistics 13: The South Carolina Meeting*. Ann Arbor, MI: Michigan Slavic Publications, 355–365.
- CARLSON, Greg (1980) *References to Kinds in English*. New York: Garland
- CHIERCHIA, Gennaro (1995) “Individual Level Predicates as Inherent Generics.” In: G. Carlson/F. Pelletier (eds), *The Generic Book*. Chicago: Chicago University Press, 176–223.
- DAHL, Östen (1970) “Some Notes on Indefinites.” *Language* 46/1, 33–41.
- DAHL, Östen (1974) “Topic-Comment Structure Revisited.” In: Ö. Dahl (ed.), *Topic and Comment, Contextual Boundedness and Focus*, Hamburg: Helmut Buske, 1–24.
- DAYAL, Veneeta (2004) “Number Marking and Indefiniteness in Kind Terms.” *Linguistics and Philosophy* 27, 393–450.

---

10 There are some other indications that subject position imposes more restrictions than object position. For instance, the non-specific marker *kakoj-nibud'* ‘some’ (cf. ex. (11c) above) cannot easily appear modifying preverbal subjects. Whether the restrictions on BSgs have anything to do with the restrictions on *-nibud'* markers remains to be established.

- DORON, Edit (2003) “Bare Singular Reference to Kinds.” In: R. B. Young/Yuping Zhou (eds), *Proceedings of SALT 13*. Seattle: University of Washington, 73–90.
- ENÇ, Murvet (1991) “The Semantics of Specificity.” *Linguistic Inquiry* 22/1, 1–25.
- ENDRISS, Cornelia (2009) *Quantificational Topics. A Scopal Treatment of Exceptional Wide Scope Phenomena*. Dordrecht: Springer.
- ERTESCHIK-SHIR, Nomi (2013) “Information Structure and (In)definiteness.” In: P. Cabredo Hofherr/A. Zribi Hertz (eds), *Crosslinguistic Studies on Noun Phrase Structure and Reference*. Leiden: Brill, 23–51.
- ESPINAL, M. Teresa/Louise MCNALLY (2011) “Bare Nominals and Incorporating Verbs in Catalan and Spanish.” *Journal of Linguistics* 47, 87–128.
- FODOR, Janet and Ivan SAG (1982) “Referential and Quantificational Indefinites.” *Linguistics and Philosophy* 5, 355–398.
- GEIST, Ljudmila (2008) “Specificity as Referential Anchoring: Evidence from Russian.” In: A. Grønn (ed.), *Proceedings of Sinn und Bedeutung* 12. Oslo: ILOS, 151–164.
- GEIST, Ljudmila (2010) “Bare Singular NPs in Argument Positions: Restrictions on Indefiniteness.” *International Review of Pragmatics* 2, 191–227.
- GEURTS, Bart (2002[2010]) “Specific Indefinites, Presupposition, and Scope.” In: R. Bäuerle/U. Reyle/T. E. Zimmermann (eds), *Presupposition and Discourse*. Bingley: Emerald, 125–158.
- HASPELMATH, Martin (1997) *Indefinite Pronouns*. Oxford: Clarendon Press.
- HEIM, Irene (1982) *The Semantics of Definite and Indefinite Noun Phrases*. PhD Dissertation. Amherst: University of Massachusetts.
- von HEUSINGER, Klaus (2011) “Specificity.” In: K. von Heusinger/C. Maienborn/P. Portner (eds), *Semantics: An International Handbook of Natural Language Meaning*. Vol 2. Berlin: de Gruyter, 1024–1057.
- IONIN, Tania (2009). “Specificity.” In: L. Cummings (ed.), *The Pragmatics Encyclopedia*. New York: Routledge.
- IONIN, Tania (2013) “Pragmatic Variation among Specificity Markers.” In: S. Hinterwimmer/C. Ebert (eds), *Different Kinds of Specificity Across Languages*. Dordrecht: Springer, 75–103.
- KARTTUNEN, Lauri (1968) “What Do Referential Indices Refer To?” *Technical Report P38543, RAND Corporation Report*. Santa Monica, California.
- KARTTUNEN, Lauri (1969[1976]) “Discourse Referents.” In: *Proceedings of the 1969 Conference on Computational Linguistics*. Sång-Säby, Sweden, 1–38. <http://portal.acm.org/citation.cfm?id=990403.990490>. [Reprinted in: J. McCawley (ed.), *Syntax and Semantics 7: Notes from the Linguistic Underground*. New York: Academic Press, 1976, 363–385.]
- KRIFKA, Manfred/Francis J. PELLETIER/Greg CARLSON/Alice ter MEULEN/Godehard LINK/Gennaro CHERCHIA (1995) “Introduction.” In: G. Carlson/F. Pelletier (eds), *The Generic Book*. Chicago: The University of Chicago Press, 1–124.
- KRATZER, Angelika (1998) “Scope or Pseudoscope? Are there Wide-Scope Indefinites?” In: S. Rothstein (ed.), *Events and Grammar*. Dordrecht: Kluwer, 163–196.

- LADUSAW, William (1994) “Thetic and Categorical, Stage and Individual, Weak and Strong.” In: M. Harvey/L. Santelmann (eds), *Proceedings of SALT IV*. Ithaca, NY: Cornell U. DMLL, 220–229.
- MUNN, Alan/Cristina SCHMITT (2005) “Number and Indefinites.” *Lingua* 115/6, 821–855.
- PERELTSVAIG, Asya (2006) “Small Nominals.” *Natural Language and Linguistic Theory* 24, 433–500.
- PORTNER, Paul/Katsuhiko YABUSHITA (1998) “The Semantics and Pragmatics of Topic Phrases.” *Linguistics & Philosophy* 21, 117–157.
- RAMCHAND, Gillian/Peter SVENONIUS (2008) “Mapping a parochial lexicon onto a Universal Semantics.” In: M. Teresa Biberauer (ed.), *The Limits of Syntactic Variation*. Amsterdam: John Benjamins, 219–245.
- REINHART, Tanya (1981) “Pragmatics and Linguistics. An Analysis of Sentence Topics.” *Philosophica* 27, 53–94.
- REINHART, Tanya (1997) “Quantifier Scope: How Labour is divided between QR and Choice Functions.” *Linguistics & Philosophy* 20, 335–397.
- STVAN, Laurel Smith (1998) *The Semantics and Pragmatics of Bare Singular Noun Phrases*. Ph.D. thesis. Chicago: Evanston.
- de SWART, Henriette (2015) “Constructions with and without Articles.” In: O. Borik/B. Gehrke (eds), *The Syntax and Semantics of Pseudo-Incorporation*. Leiden: Brill. 126–156.
- ŠVEDOVA, Natalija Ju. (ed.), (1982) *Russkaja grammatika*. Moscow: Nauka.
- WINTER, Yoad (1997) “Choice Functions and the Scopal Semantics of Indefinites.” *Linguistics & Philosophy* 20, 399–467.
- YANOVICH, Igor (2005) “Choice-Functional Series of Indefinites and Hamblin Semantics.” In: E. Georgala/J. Howell (eds), *Proceedings of Semantics and Linguistic Theory 15*. Ithaca, NY: Cornell University, 309–326.



Summary  
CONSTRAINTS ON THE POSITION AND INTERPRETATION  
OF BARE SINGULAR INDEFINITES IN RUSSIAN

In this paper, I discuss the distribution of bare singular indefinite nominals in Russian. I argue that these nominal phrases are full-fledged indefinites and can have both specific and non-specific indefinite interpretations. However, their appearance in the preverbal subject position is not common. I argue in this paper that this restriction exists because a specific interpretation, or, in other words, a referential reading, required for this position cannot be unambiguously established with BSGs without any additional sentential specification. Additionally, I discuss some consequences of my proposal, such as BSG subjects of individual-level predicates, and BSG subjects inthetic judgments.

**Keywords:** indefinite, Russian, specificity, topic

Povzetek  
OMEJITVE POZICIJE TER INTERPRETACIJE GOLIH SAMOSTALNIKOV  
V EDNINSKI NEDOLOČNI OBLIKI V RUŠČINI

Članek obravnava distribucijo golih samostalnikov v edninski nedoločni obliki v ruščini. Zagovarja trditev, da so te samostalniške zveze polno razvite nedoločniške oblike, ki lahko prejmejo tako specifično kot nespecifično nedoločno interpretacijo. Vseeno pa se redko pojavljajo pred glagolom v položaju osebka. V članku je ta omejitev analizirana kot posledica dejstva, da specifično oz. nanosniško branje, ki ga zahteva ta položaj, pri golih samostalnikih v edninski nedoločni obliki ne more biti nedvoumno vzpostavljeno brez dodatne stavčne specifikacije. Članek obravnava tudi nekaj posledic, ki sledijo iz predlagane analize.

**Ključne besede:** nedoločna oblika, samostalnik, ruščina, specifičnost, izhodišče







## WHAT IS SENT TO SPELL-OUT IS PHASES, NOT PHASAL COMPLEMENTS\*\*

### 1. INTRODUCTION

An appealing property of the phase theory, emphasized already in Chomsky (2000), is that phases are relevant to many phenomena. In fact, given the variety of the phenomena where phases have been argued to be relevant, we may expect all domain-based mechanisms to be stated in terms of phases, with phases defining their locality domain.

One of the mechanisms where phases are crucially assumed to be involved is multiple spell-out. The standard assumption in the Government and Binding framework as well as early minimalism was that there is a single point of spell-out, where the derivation separates into the phonology branch and the semantics branch. The advent of minimalism, where one of the guiding hypotheses is that UG is an optimal way of satisfying requirements imposed on the language faculty by the external systems that the language faculty interfaces with, led to the elimination of DS and SS, the “internal” levels of representation which do not interface with the external systems. While under Chomsky’s (1993) single spell-out approach PF and LF considerations determine when spell-out applies,<sup>1</sup> the assumption that spell-out applies only once, which was adopted in that work, still hid a trace of SS. A number of authors (see especially Uriagereka 1999; Epstein 1999; Epstein et al 1998; and Chomsky 2000, 2001) have argued that Spell-Out may apply multiple times, which has led to a radically derivational nature of the computation, with the interfaces accessing syntactic computation as the derivation proceeds without the mediation of PF and LF levels of representation.

A question that arises under the multiple spell-out approach is what determines spell-out units. The standard assumption is that phases do that. In the original proposal regarding phases (Chomsky 2000), what was sent to spell-out was the phase itself (see also Franks and Bošković 2001). The assumption was later modified in that what was sent to spell-out was assumed to be the complement of the phase head (referred to below as phasal complement for ease of exposition), which is in fact the standard

---

\* zeljko.boskovic@uconn.edu

\*\* The article is based upon work supported by the NSF under Grant BCS-0920888. For helpful suggestions, I thank Renato Lacerda, Adrian Stegovec, two anonymous reviewers, the participants of my Fall 2016 UConn seminar, and the audiences at SinFonJA 8, Tsinghua University (Beijing), University of Sarajevo, and University of Pennsylvania.

1 Movement that takes place prior to spell-out was assumed to be driven by the need to eliminate strong features, which are illegitimate PF objects. Procrastinate, which favors LF movement, was assumed to delay movement that is not motivated by PF considerations until LF.

assumption currently.<sup>2</sup> This has, however, left us with a rather strange situation: while phases are the crucial units in the multiple spell-out framework, for all practical purposes the crucial units are actually not phases but phasal complements. But, in contrast to phases, phasal complements have no theoretical status. Witness e.g. the great deal of effort that has gone into coming up with a proper, unified definition of what counts as a phase; there has been nothing like that regarding phasal complements—no one has even tried to come up with a unified definition of phasal complements. The reason for that is quite obvious: phasal complements have no theoretical status in phase theory. But phasal complements should then play no role in spell-out; what is transferred to spell-out should be phases, not phasal complements. The goal of this article is to explore the viability of such an approach to spell-out, which will in fact be argued to be superior in several respects to the phasal complement spell-out approach. As we will see below, the real reason for the modification of the original approach in this respect, i.e. for adopting the assumption that what is sent to spell-out is the phasal complement, not the phase itself, is actually successive-cyclic movement, more precisely, the interaction of successive-cyclic movement and multiple spell-out. As a result, arguing for phasal spell-out will inevitably involve re-examining successive-cyclic movement. Most of the article will therefore be devoted to exploring the interaction of multiple spell-out and successive-cyclic movement.

I will start the discussion by noting in section 2 a rather straightforward argument for phasal spell-out regarding matrix clauses. In section 3 I show that Chomsky's (2013) labeling system also favors phasal spell-out. Section 4 presents several arguments for phasal spell-out regarding syntax-phonology interaction, in particular, cliticization in Bulgarian and Arabic, stress assignment in German, raddoppiamento fonosintattico in Abruzzese, and tone sandhi in Taiwanese. In section 5 I turn to the interaction of successive-cyclic movement and multiple spell-out.<sup>3</sup> The adoption of phasal spell-out will lead us to an approach to successive-cyclic movement which is rather different from the one that is currently standardly assumed, the reason being that the latter looks the way it does because of the assumption that what is sent to spell-out is the phasal complement, not the phase itself. Since the analysis from section 5 has a number of consequences, the bulk of the article will be devoted to the issues discussed in that section (the main phenomena discussed in that section concern locality of movement and stress assignment), some of which go beyond the main goal of this article, which is to make a case for phasal spell-out, but are necessary since they are affected by the phasal spell-out analysis of successive-cyclic movement. Finally, section 6 concludes the article.

---

2 Chomsky (2000:131-132) suggests that phases are sent to spell-out. Chomsky (2001) is already somewhat ambivalent in this respect: thus, Chomsky (2001:12) suggests phasal spell-out, while Chomsky (2001:13) suggests phasal complement spell-out.

3 The discussion in that section is based on Bošković (2015b), putting the analysis from that work in a broader perspective (i.e. framing it within more general theoretical concerns), and also adding new arguments.

## 2. MATRIX CLAUSES

In addition to the obvious conceptual argument which concerns the fact that phasal complements have no theoretical status in phase theory (only phases do), a rather straightforward argument for the original assumption that what is sent to spell-out is the phase itself concerns matrix clauses. The phasal spell-out approach does not require special provisos regarding matrix clauses. They, however, significantly complicate the phasal complement spell-out approach. In particular, if only phasal complements are sent to spell-out, matrix clauses will never be sent to spell-out. Under this approach, we then need an additional assumption regarding matrix clauses to ensure that not just IP but also CP, which is a phase, not a phasal complement, is sent to spell-out in (1).

(1) [<sub>CP</sub> What did [<sub>IP</sub> John buy]]?

In other words, under the phasal complement spell-out approach we still need to assume that in certain cases (i.e. matrix clauses) what is sent to spell-out is a phase. This is not an innocent assumption given that in many languages, matrix and embedded clauses look exactly the same, which means that at the point where the clause itself is built we cannot tell whether we are dealing with an embedded or a matrix clause. This is e.g. the case with the embedded and the matrix clause in English (2a) and (2b) and the embedded and the matrix clause in English (3a) and (3b), though it is easy to come up with even more straightforward examples of this sort for many languages.<sup>4</sup>

(2) a) *John left.*                      b) *I think John left.*  
(3) a) *Who left?*                      b) *I wonder who left.*

The assumption that what is sent to spell-out is phasal complements thus requires an additional non-innocent proviso to handle matrix clauses, thus also requiring phasal spell-out in this case. On the other hand, the phasal spell-out approach is consistent in that only phases are ever sent to spell-out under this approach.

## 3. LABELING AND PHASAL SPELL-OUT

One of the issues that was taken for granted in the GB framework but has undergone close scrutiny within minimalism concerns labeling of the structure created in the syntax, i.e. determining which element projects when two elements undergo merger. To take a concrete example, it was standardly assumed in early minimalism that when e.g. *arrive* and *John* are merged {*arrive*, *John*}, information needs to be provided that the resulting object is verbal in nature. In more technical terms, *arrive* projects, labeling the resulting object. In early minimalism labeling was taken to be part of the definition

---

4 See Agbayani (2000), An (2007), Boeckx (2003), Bošković (2016b), McCloskey (2000), Petsky and Torrego (2001), among others, for arguments that examples like (3a-b) involve movement of *who* to SpecCP.

of Merge: when X and Y are merged, either X or Y projects, i.e. labels the resulting object. Collins (2002), however, questioned the need for labeling, arguing that labeling should be eliminated. Chomsky (2013, 2015) has recently proposed a system that is in between early Minimalism and Collins (2002) in this respect, where unlabeled objects are allowed during the derivation but not in the final representation. In particular, like Collins, Chomsky (2013, 2015) argues that labeling is not part of the definition of Merge, Merge being a set-formation operation that combines two objects without label projection. However, Chomsky keeps the traditional assumption that when X and Y are merged, the nature of the resulting object needs to be specified. In the case at hand, it still needs to be specified that the object formed by the merger of *arrive* and *John* is verbal in nature. Crucially, Chomsky argues that syntax itself does not require this information; the information is required by the interfaces so that syntactic objects can be interpreted. In other words, although there is nothing wrong with unlabeled objects in the syntax, such objects are uninterpretable. Chomsky then provides an algorithm that specifies labels which applies at the point of transfer to the interfaces, labeling being interface-driven.

Under Chomsky's interpretation-driven approach to labeling we would expect labeling to take place at the point of transfer to the interfaces for what is being transferred to the interfaces. Under Chomsky's approach to spell-out, this should be the phasal complement. This is, however, not what Chomsky assumes. Although Chomsky assumes that what is sent to spell-out is the phasal complement, he crucially assumes that the whole phase is labeled at this point, not just the phasal complement. There is a clear inconsistency here; if labeling is strictly interpretation/interface-driven, with syntax itself not requiring labeling, what should be labeled is what is sent to the interfaces. The issue is, however, resolved if what is sent to the interfaces is indeed the whole phase. It is then not surprising that what is labeled is the whole phase, not just the phasal complement. Chomsky's (2013, 2015) labeling system, where labeling is assumed to be interface-driven and what is being labeled is phases, thus also naturally leads to the assumption that full phases, not phasal complements, are sent to spell-out.

#### **4. ARGUMENTS FOR PHASAL SPELL-OUT FROM SYNTAX-PHONOLOGY INTERACTION**

In this section I will present several arguments for phasal spell-out regarding syntax-phonology interaction.<sup>5</sup> The arguments come from cliticization in Bulgarian, Macedonian and Arabic, stress assignment in German, raddoppiamento fonosintattico in Abruzzese, and tone sandhi in Taiwanese.

Some of the original arguments for multiple spell-out were in fact based on the assumption that what is sent to spell-out is the phase itself. I will start the discussion by

---

5 A potentially interfering factor here is the possibility of prosodic factors influencing spell-out domains (which can be implemented through readjustments of the initial prosodic phrasing), an issue that cannot be addressed here; for a recent survey of the relevant literature see Cheng and Downing (2016).

summarizing one such argument from Franks and Bošković (2001), which concerns cliticization in Bulgarian and Macedonian.<sup>6</sup> (4) illustrates the basic cliticization pattern for these languages (clitics are in boldface).

- |        |                                      |        |         |
|--------|--------------------------------------|--------|---------|
| (4) a) | <i>Vera <b>mi go</b> dade včera</i>  | Bg: OK | Mac: OK |
|        | Vera me.dat it.acc gave yesterday    |        |         |
|        | ‘Vera gave me it yesterday.’         |        |         |
| b)     | <i>Vera <b>mi go</b> včera dade.</i> | Bg: *  | Mac: *  |
| c)     | <i><b>Mi go</b> dade Vera včera.</i> | Bg: *  | Mac: OK |
| d)     | <i>Dade <b>mi go</b> Vera včera.</i> | Bg: OK | Mac: *  |

While in Macedonian the clitics always precede the verb, in Bulgarian they precede the verb except when preceding the verb would leave them in sentence initial position; in that case they follow the verb. What is going on here is rather straightforward. Bulgarian clitics are enclitics, i.e. they must encliticize, while Macedonian clitics procliticize. The V-clitic order is a last-resort strategy employed in Bulgarian when the enclitic requirement otherwise cannot be satisfied.

Bošković (2001) proposes a copy-and-delete account of the effect in question, based on a proposal regarding pronunciation of lower copies made in Franks (1998), which was later argued for by many authors. In particular, the proposal is that, as in the case of deletion of copies applying in LF, with PF deletion of copies, there is no rigid requirement to always pronounce one particular copy: a lower copy of a non-trivial chain can be pronounced in PF iff this is necessary to avoid a PF violation (see Abels 2001; Bobaljik 2002; Bošković 2001, 2002a; Bošković and Nunes 2007; Franks 1998; Lambova 2002; Landau 2003; Pesetsky 1998; Reglero 2007; Stjepanović 2003, among others, for many cases of such lower copy pronunciation motivated by PF considerations.)

- |        |             |       |
|--------|-------------|-------|
| (5) a) | X cl V eł   | (Bg)  |
| b)     | eł V cl     |       |
| c)     | (X) cl V eł | (Mac) |

In both Bulgarian and Macedonian the clitic undergoes movement, so that there is a copy of the clitic both preceding and following the verb. In Macedonian, where clitics are proclitics, nothing goes wrong if the higher copy of clitic movement is pronounced, hence this copy must be pronounced, the lower copy being deleted in (5c). As a result, the cl-V order always obtains. When a phonologically realized element precedes a clitic in Bulgarian, as in (5a), the clitic can encliticize; since nothing goes wrong if the higher copy of the clitic is pronounced in (5a) this copy must be pronounced. However,

6 Another relevant work based on phasal spell-out is Fox and Pesetsky (2005). Since the Fox and Pesetsky argument is a bit more involved I refer the reader to the original work for discussion. It should, however, be noted that they do not share the assumption adopted in e.g. Chomsky (2000, 2001) and Uriagereka (1999), also adopted in this article, that the internal structure of what is sent to spell-out is inaccessible to the syntax.

pronunciation of the higher copy would lead to a PF violation in (5b) since the enclitic requirement would not be satisfied. This is precisely the case where lower copy pronunciation is allowed, as in (5b). The analysis captures the above paradigm, including the last resort nature of the V-clitic order in Bulgarian, without positing any syntactic differences between clitics in Bulgarian and Macedonian, which is desirable given that syntactically, the clitics in this type of examples in Bulgarian and Macedonian behave in the same way in all respects.

Consider now (6), from Franks and Bošković (2001).

- (6) a) *I ti go dade.* (Bg)  
 and you.dat it.acc gave  
 ‘And (s)he gave it to you.’  
 b) \**I dade ti go.*

What these data indicate is that clitics can encliticize to the conjunction. As a result, we get the clitic-V order here, V-clitic order being disallowed.

Consider now the following paradigm from Franks and Bošković (2001).

- (7) a) *I ti go dade li?*  
 and you.dat it.acc gave Q  
 ‘And did (s)he give it to you?’ (Mac)  
 b) *I dade li ti go?*  
 and gave Q you.dat it.acc  
 ‘And did (s)he give it to you?’ (Bg)  
 c) \**I ti go dade li?* (Bg)  
 d) \**I ti go li dade?* (Bg)  
 e) \**I li ti go dade?*  
 ‘And did (s)he give it to you?’ (Bg)

The standard account of Bulgarian and Macedonian clitics is that they are adjoined to the verb (see e.g. Bošković 2001, 2002b; Franks and King 2000; note in this respect the ungrammaticality of (4b)). As a result, when the verb undergoes head-movement it carries the clitics with it, as illustrated by Macedonian (7a). The question particle *li* triggers V-to-C movement in Macedonian and Bulgarian. V-movement carries the clitics along, hence the complex clitics+V precedes the complementizer *li* in (7a), which is the only possible word order here in Macedonian.

An unexpected word order, however, surfaces in Bulgarian. Recall the V-clitic order in Bulgarian is obtained through last-resort deletion of the highest clitic copy, when this is necessary to satisfy the enclitic requirement. Recall also that the conjunction can satisfy it, as shown by (6). One would then expect (7a) to be acceptable in Bulgarian too, given that the word order difference between the two languages noted above (regarding the V-clitic order) arises only when there is nothing else in front of the clitics in Bulgarian. However, the word order of (7a) is unacceptable in Bulgarian, as (7c)

shows. Instead, what we get is (7b), all other orders being unacceptable (7d-e). This is rather surprising. Why is it that lower copy pronunciation applies in (7b), although it does not in (6a)? Franks and Bošković (2001) provide a multiple spell-out account of these data, which is however crucially based on the original assumption that what is sent to spell-out is the phase itself.

Following Chomsky (2000), Franks and Bošković (2001) assume that CPs, but not IPs, are phases. Given that phases, not phasal complements, are sent to spell-out, this means that CPs, but not IPs, are sent to the phonology. Franks and Bošković also argue that the conjunct following *i* is a CP in (7), since it is headed by *li*, a C element. However, it is a bare IP in (6). Given that the conjunct in (6) is an IP, it will not be sent to the phonology until the whole structure is built.<sup>7</sup> Since clitics can encliticize to *i* all elements can then be pronounced in the highest position in (6).

On the other hand, in the *li* construction spell-out applies as soon as the conjunct following *i*, which is a CP hence a phase, is built. To avoid having a stranded enclitic in PF, *ti go* are pronounced in a lower position.

(8) [<sub>CP</sub> [<sub>C</sub> ~~ti go~~ dade+li] ti go dade]

The conjunction *i* is only then added and we derive (7b):

(9) i [<sub>CP</sub> [<sub>C</sub> ~~ti go~~ dade+li] ti go dade]

As for Macedonian, since its pronominal clitics can be proclitics, higher copies can be pronounced even without *i*.

(10) a) [<sub>CP</sub> [[ti go dade] li] ~~mi go dade~~  
 b) i [<sub>CP</sub> [[ti go dade] li] ~~mi go dade~~]

The paradigm in question provides an argument for multiple spell-out. The issue here is that the conjunction in Bulgarian (7) behaves as if it were not there (in contrast to (6)). The best and the most straightforward way of capturing this kind of situation where an element behaves as if it were not there is that it is indeed not there. Under multiple spell-out, the conjunction is actually not there at the relevant point of the derivation, which straightforwardly captures its invisibility. Notice, however, that the crucial component of the analysis is that what is sent to spell-out is the phase itself, not its complement. If phasal complements were to be sent to spell-out, what would be sent to spell-out in (7) would be the IP, which means that the clitics-verb complex, which precedes the complementizer, would not be sent to spell-out before the conjunction is merged into the structure, incorrectly predicting (7c) to be acceptable and (7b) to be unacceptable, on a par with (6). The analysis thus provides an argument for phasal spell-out.

7 The implicit assumption here is that ConjPs are phases. (In fact, with an additional assumption (namely, that merger with a projection of Conj<sup>0</sup> also triggers spell-out), the locality system from section 5 can be extended to the Coordinate Structure Constraint if ConjP is a phase.)



There are other PF phenomena which also indicate that there is a spell-out boundary at the phasal edge. Consider stress assignment in all-new neutral sentences in German. The phenomenon has been discussed extensively in the literature. It would be way beyond the scope of this article to discuss it in any detail here. I will focus on a couple of constructions, which will be analyzed within the overall approach to phase-based pitch accent assignment from Kratzer and Selkirk (2007), where pitch accent assignment proceeds cyclically, with phases (i.e. spell-out) determining its domain.<sup>8</sup> The relevant assumptions from Kratzer and Selkirk (2007) are given below (slightly modified):

- (11) The spell-out domain is the prosodic domain for phrase stress.
- (12) Phrase stress is assigned within the highest phrase within the spell-out domain.
- (13) The Elsewhere Condition on Prosodic Spell-out: Phrase stress must be assigned within a spell-out domain with eligible material.

Consider in this respect (14), assuming an all-new neutral context. Kratzer and Selkirk argue that the bracketed element corresponds to a spell-out domain. What is important to note here is that the highest phrase within the domain receives phrase stress (*eine Géige*).

- (14) ... *dass ein Jünge [eine Géige an einen Freund schickte]*.  
           that a boy a violin to a.acc friend sent  
           ‘...that a boy sent a violin to a friend.’ (Kratzer and Selkirk 2007:107)

The question to address now is whether the spell-out domain corresponds to a full phase or a phasal complement. There is reason to believe that it is the former. Consider (15), where both the subject and the object of the embedded clause receive phrase stress.

- (15) *Ich glaube, dass Maria die Gesetze studiert.*  
       I think that Maria the laws is.studying  
       ‘I think that Maria is studying the laws.’ (Kratzer and Selkirk 2007:94)

The line of research going back to Kayne (1994) and Zwart (1993) argues that the OV order of languages like German results from object shift, i.e. movement to SpecvP. Even independently of this line of research, Diesing (1996) argues that non-contrastive definite objects must undergo movement in German. In other words, the object in (15) is generally assumed to undergo object shift/scrambling.

Given antilocality (the ban on movement that is too short), which bans movement within a phrase (see Bošković 1994, 2013a; Saito and Murasugi 1999; Abels 2003; Grohmann 2003, among many others), the movement in question can target SpecvP or a higher position, i.e. SpecvP is the lowest potential landing site for this movement.

---

8 I will, however, depart from Kratzer and Selkirk in some aspects of the syntax of the relevant constructions (see also Kahnemuyipour 2003 for an early version of (12)).

Now, for the object to be stressed it cannot be located in the same spell-out domain as the subject. If the object moves to SpecvP, and what is sent to spell-out is the whole vP phase, the object and the subject will be in different spell-out domains, so that each element will receive stress. On the other hand, if what is sent to spell-out is the phasal complement, which is VP, the object and the subject would end up in the same spell-out domain, which means the object could not be stressed. The phasal spell-out system thus captures the stress pattern in (14), where both the subject and the object bear phrase stress, which is problematic for the phasal complement spell-out system.

Consider now the contrast between (16) and (17).

- (16) *Ich hab' gelesen, dass die Metallarbeiter gestreikt haben.*  
 I have read that the metal.workers gone.on.strike have  
 'I read that the metal workers went on strike.' (Kratzer and Selkirk 2007:95)
- (17) *Ich hab' gehört, dass der Rhén stinkt.*  
 I have heard that the Rhine stinks  
 'I've heard that the Rhine stinks.' (Kratzer and Selkirk 2007:96)

(16) involves a stage level predicate, and (17) an individual level predicate. Diesing (1992) argues that the subject in (16) can be located either in SpecvP or SpecTP in overt syntax, while the subject in (17) must be located in SpecTP. Assume this to be the case.<sup>9</sup> What is important here is that the verb can be unstressed only in (16). Consider how these data can be handled under phasal spell-out. For (16), a derivation is available where the subject and the verb are in the same spell-out domain, the subject staying in SpecvP. On this derivation, the subject receives phrase stress as the highest phrase within the vP spell-out domain, which corresponds to the whole vP phase. This is not an option in (17), where the subject cannot be in SpecvP. Since the subject is located outside of vP in (17), the verb, which I assume is located in v, is the only element within the vP phase, which is sent to spell-out, hence it receives stress under (13). ((13) assigns stress to the verb, which is a head, not a phrase, when there is nothing else in the spell-out domain that can receive stress.)<sup>10</sup> Not stressing the verb is thus not an option for (17) under phasal spell-out. This is, however, not the case under phasal complement spell-out. Without adopting any additional assumptions, the verb, located in v, should not get stress in (17) since it belongs to the same spell-out domain as the subject in

---

9 It should be noted that the issue in question is not settled. Assuming split IP, where IP splits into more than one phrase, makes the issue in question even trickier. What is needed for the argument given below to go through under such clausal structure is merely that subjects of individual level predicates cannot be located in the  $\theta$ -position in overt syntax (in fact, simply assuming that subjects of stage level predicates can be lower than subjects of individual level predicates may be enough; see in this respect the discussion of *was-für* split and the split-topic construction in Diesing 1992).

10 Since the subject of stage-level predicates can move outside of vP we would expect that the verb in examples like (16) can be optionally stressed. According to Kratzer and Selkirk (2007), this is indeed the case.

spite of the subject being located outside of vP, given that what is sent to spell-out is VP. The no stress on the verb should in fact be the only option for both (17) and (16) since regardless of whether the subject moves to SpecTP or not it would be in the same spell-out domain as the verb. The facts regarding stress assignment in all-new neutral sentences in German thus favor phasal spell-out.<sup>11</sup>

The arguments for full phase spell-out from Bulgarian cliticization and stress assignment in German came from the impossibility of interaction between the edge of phase XP and material outside of phase XP, which is predicted under phasal spell-out given that the two belong to different spell-out domains, in contrast to phasal complement spell-out, where the two are in the same spell-out domain. A relevant argument can in principle also be constructed by looking at potential PF interaction between a phasal head and its complement. Under phasal spell-out, such interaction should be possible (for phase-sensitive phenomena), since the two belong to the same spell-out domain. On the other hand, such interaction should be impossible under phasal complement spell-out since the two do not belong to the same spell-out domain. There are several phenomena which indicate that such interaction is possible. One case of this

11 Ishihara (2007) provides an argument for phasal spell-out based on Miyawaga's (2003) account of scopal interaction in Japanese (i)-(ii) (but see Saito 2009).

- (i) *zen'in-ga sono tesuto-o uke-nakat-ta.* (SOV)  
 all-nom that test-acc take-neg-pst  
 'All did not take that test.' \*not >> all, all >> not (Ishihara 2007:139)
- (ii) *sono tesuto-o<sub>i</sub> zen'in-ga t<sub>i</sub> uke-nakat-ta.* (OSV)  
 that test-acc all-nom take-neg-pst  
 'That test, all didn't take.' not >> all, all >> not (Miyagawa 2003:183-184)

Assuming SpecTP must be filled in Japanese, Miyagawa argues that the subject that moves to SpecTP must scope over the negation, while the negation obligatorily scopes over the subject that stays in its base position in SpecvP. The subject in (i) moves to SpecTP, hence it scopes over the negation. Miyawaga argues that (ii) is structurally ambiguous. On one derivation the subject moves to SpecTP, with the object undergoing A'-movement to SpecCP. On this derivation, the subject takes wide scope. On the reading where the negation takes wide scope, the object undergoes A-movement to SpecTP, which satisfies the EPP, with the subject staying in its base position. Ishihara (2007) observes that prosody disambiguates the two readings that (ii) has. When there is a major phrase boundary after the subject, the subject must take wide scope (see (iii)), while under the prosodic phrasing in (iv), the subject cannot take wide scope. Furthermore, the unambiguous (i), where the subject moves to SpecTP, patterns with (iii) in this respect, as (v) shows.

- (iii) (all >> not) [<sub>CP</sub> sono tesuto-o<sub>i</sub> [<sub>TP</sub> zen'in-ga<sub>j</sub> [<sub>VP</sub> t<sub>i</sub> t<sub>j</sub> [<sub>VP</sub> t<sub>i</sub> uke-nakat-ta]]]]  
 ( sono tesuto-o )<sub>MaP</sub> ( zen'in-ga )<sub>MaP</sub> ( uke-nakat-ta )<sub>MaP</sub>
- (iv) (not >> all) [<sub>TP</sub> sono tesuto-o<sub>i</sub> [<sub>VP</sub> t<sub>i</sub> zen'in-ga [<sub>VP</sub> t<sub>i</sub> uke-nakat-ta]]]  
 ( sono tesuto-o )<sub>MaP</sub> ( zen'in-ga uke-nakat-ta )<sub>MaP</sub>
- (v) (all >> not) [<sub>TP</sub> zen'in-ga<sub>i</sub> [<sub>VP</sub> t<sub>i</sub> sono tesuto-o uke-nakat-ta]]  
 ( zen'in-ga )<sub>MaP</sub> ( sono tesuto-o uke-nakat-ta )<sub>MaP</sub> (Ishihara 2007:147)

To account for these facts, Ishihara proposes that major phrases correspond to spell-out domains and that spell-out domains crucially correspond to full phases, not phasal complements. As a result, vP, not its VP complement, is sent to spell-out, hence the major phrase boundary at the edge of vP in (iii)-(v) (see Ishihara's work for details of the analysis).

sort is raddoppiamento fonosintattico (RF) in Abruzzese, discussed in D'Alessandro and Scheer (2015), following Biberauer and D'Alessandro (2006), D'Alessandro and Roberts (2010). In RF, the initial consonant of a word undergoes gemination which is conditioned by the properties of the preceding word. In Abruzzese, a lexically conditioned set of words triggers RF when they are in a specific syntactic relationship with the following word. Importantly, the complementizer *chə* is an RF trigger which triggers RF on the first word within its IP complement, as shown in (18). Another relevant case is provided by (19), where RF applies between the relative pronoun in SpecCP and the subject. If RF is spell-out domain sensitive (i.e. if it cannot take place between elements in different spell-out domains), we then have evidence here that phasal complements are not spell-out domains.

- (18) a) *Jè mmeje chə vve.*  
 is better that come.3sg  
 'It's better that he/she comes.'  
 b) *Jè mmeje chə nni vve.*  
 is better that not come.3sg  
 'It's better that he/she doesn't come.' (D'Alessandro and Scheer 2015:614)
- (19) *lu waglionə chə tu si vistə*  
 the boy whom you are seen  
 'the boy whom you saw' (D'Alessandro and Scheer 2015:615)

D'Alessandro and Scheer (2015) do provide evidence that RF is spell-out domain sensitive based on the impossibility of RF between the auxiliary and the participle in (20) (such RF is in principle possible, see passive (21)), which they analyze in terms of sensitivity of RF to spell-out domains. From the current perspective, (20) is analyzed as having the auxiliary outside vP, or, more precisely, the phasal domain projected by the verb (see section 5). Since the auxiliary and the participle are then located in different phases they belong to different spell-out domains (given that phases correspond to spell-out domains), hence the impossibility of RF between the auxiliary and the participle in (20).

- (20) *So rəspəttatə la leggə.*  
 am.1sg respected.sg the.f.sg law.f.sg  
 'I have respected the law.' (D'Alessandro and Scheer 2015:611)

While RF is blocked between the auxiliary and the participle with transitives, it is not blocked with passives (21), which is not surprising. Interestingly, it is blocked with ergatives, as in (22).

- (21) *So rrəspəttatə (da tuttə quində).*  
 am.1sg respected.sg by all  
 'I am respected by everybody.' (D'Alessandro and Scheer 2015:612)

- (22) *so rəmastə.*  
 am.1sg stayed.sg  
 ‘I have stayed.’

(D’Alessandro and Scheer 2015:613)

This pattern can be captured under full phase spell-out in the phasal system adopted below, where the phasal domain corresponds to the thematic domain. In that system, there is a phase with both unergative and ergative verbs. From this perspective, RF in passives can be captured if the auxiliary in the passive case is located in the voice head, which belongs to the thematic domain due to the absorption of the external  $\theta$ -role (the highest projection in the verbal thematic domain, which is actually different in the structures below, functions as a phase, see section 5).<sup>12</sup>

- (23) a) auxiliary [<sub>verbal thematic domain-VP</sub> participle] active unergative  
 b) auxiliary [<sub>verbal thematic domain-VP</sub> participle] ergative  
 c) [<sub>verbal thematic domain-VoiceP</sub> auxiliary participle] passive

The brackets indicate the relevant phase in (23). Since there is a phasal boundary between the auxiliary and the participle with active unergatives and ergatives but not with passives, the RF in question is possible only with passives.

As noted by an anonymous reviewer, in principle similar arguments can be provided by languages where subject clitics cliticize to a complementizer. Adrian Stegovec (p.c.) notes a case of this sort from Standard Arabic,<sup>13</sup> which involves phonological interaction between complementizers and subject clitics (the subject in the relevant constructions must be accusative; for general discussion of clitics in Arabic, see Shlonsky 1994, 1997 and Fassi Fehri 1993, among others; on complementizers in Arabic, see these works as well as Shlonsky 2000; Fassi Fehri 1988; Khalaily 1995; and Benmamoun 2000, among others). The relevant clitics start with a consonant, and the complementizer that hosts them ends with a vowel (e.g. *ʔanna*). The only exceptional clitic in this respect is the 1p.sg one, whose exact form depends on its host. Albuhayri (2013) argues that the form of the 1p.sg ACC clitic is determined solely by the phonological properties of the host.<sup>14</sup> The basic form of the clitic is *ya*, which is part of the strong

12 The 3p.sg form of BE, *jè*, triggers RF on C in (i). This may not be surprising in the phasal system from Bošković (2015a, 2016b), where the highest projection in the thematic domain of a lexical head and the highest projection in the non-thematic domain function as phases. Under this conception, *jè* actually may not start a new lexical head thematic phasal domain, which means CP would not function as a phase here (the phasehood would extend all the way to the matrix CP), hence RF is possible.

(i) *Chi jè cchə l’a fattə?*  
 who is that it-has done  
 ‘Who is it that did that?’ (Roberta D’Alessandro (p.c.))

13 The discussion of Arabic below is based on an analysis suggested by Adrian Stegovec (p.c.).

14 Salih (1985) and Holes (2004) argue that the clitic form depends on the category of the host. Salih suggests *ii* occurs with prepositions and *nii* with other hosts, while Holes suggests *nii* is used with verbs and *ii* with nouns and prepositions. The statements are, however, incorrect. Thus, *nii* can

pronominal form that marks number and person information (unless otherwise noted, the data are taken from Albuhayri 2013).

- (24) *ʔiyyaa-ya* = “me” (-*ya* = 1sg; *ʔiyyaa* = support morpheme shared by all accusative strong pronouns)

Two other forms, *ii* and *nii*, are allomorphs of *ya*. *Ya* occurs when the host ends in a long vowel:

- (25) *ʃaʃaa-ya* = stick.1sg = “my stick”

*Ya* optionally alternates with *ii* when attached to a noun that ends in a consonant (which is not *n*):

- (26) a) *kitaab-ii* -> syllabified as [ki.taa.bii] = “my book”  
 b) *kitaab-ya* -> syllabified as [ki.taab.ya] = “my book”

When *nii* surfaces, *n* is used epenthetically. This e.g. happens with verbs with a vowel final suffix:

- (27) *qaabal-a-n-ii*  
 met-perf.3m.sg-/n/-1sg  
 ‘He met me.’

Clitic pronouns must have an initial onset (they cannot be \*V(V), only CV(V)), see in this respect the syllabification in (26a)), which can in principle be satisfied in two ways in cases like (27): by deleting the short vowel of the host, or by inserting *n*. The first option is not available in (27), because deleting the vowel would make the perfective subject agreement marker unrecoverable, so the *n*-insertion strategy is used. On the other hand, with case marked nouns with a 1p.sg ACC possessive clitic, the opposite strategy is used: since the case information is recoverable the case suffix is deleted and the clitic surfaces as *ii* or *iya* (with an epenthetic *i* vowel).<sup>15</sup>

---

occur with complementizers (28) and prepositions as well; compare in this respect (i) and (ii) (the issue here is that the final *n* of the preposition cannot be resyllabified into the onset of the clitic; i.e., the alternation is phonologically conditioned, as discussed below).

- (i) a. *lii*/\**li-nii* (li=to)  
 to-me  
 b. *min-nii* (min=with)  
 with-me

Note also that, as expected given the discussion below, prepositions ending in a long final vowel require the *ya* clitic form ((ii) involves diphthongization of the final vowel of the preposition).

- (ii) *fiy-ya* (fi=in, at)

15 Consider in this respect (i) (Albuhayri 2013:49); note that the nominative form would have the *-u* ending, as in *qalam-u* (the accusative is *qalam-a* and genitive *qalam-i*).

The main point here is that all clitic forms can be determined from phonological information and the different strategies that can be employed so that the host+clitic sequence fits one of the syllabification templates of Arabic. Albuhayri (2013) does not discuss the relevant complementizer+clitic constructions, but Adrian Stegovec (p.c.) observes that there is phonological interaction between the two. The complementizers that trigger subject cliticization are *ʔinna* and *ʔanna*. The clitic form that is used with the complementizers is *nii*, as in (28).<sup>16</sup>

- (28) qaal-at ʔinna-**nii** ʔahab-tu ʔila l-qaryat-i  
 said-3f.sg that-1sg.acc went-1sg to the-village-obl  
 ‘She said that I went to the village.’ (Salih 1985:42)

What is of interest here is that the complementizer is neither VV-final nor C-final. The relevant complementizers are morphologically complex. *ʔanna* has a counterpart that does not co-occur with a clitic, *ʔan*. Khalaily (1995) in fact analyzes the complementizer as being composed of *ʔan* and *na*. As for *ʔinna*, Benmamoun (2000) notes the *-na* segment is a focus particle (*ʔinna* is an emphatic complementizer), a treatment that can also be applied to *ʔanna*. Adrian Stegovec (p.c.) observes that, given this, the *-na* part of these complementizers cannot be deleted when the 1p.sg clitic is attached to make the host+clitic fit a syllabic template. This is because *na* is a separate morpheme, which cannot be recovered upon deletion. Instead, the epenthetic *n* is inserted, so the clitic surfaces as *nii*. What is important here is that the form of the complementizer affects the PF realization of the clitic.

It should be noted that Palestinian Arabic may be different in this respect. Thus, Shlonsky (1994:2) gives examples where the form of the complementizer changes when a clitic is attached:

- (29) a) *ʔinnu l-mħalme*                      b) *ʔin-ha*  
 that the-teacher                              that-she

The *ʔinnu* complementizer which otherwise does not appear in the short form surfaces as *ʔin* when a subject accusative clitic is attached. (29) from Palestinian Arabic is then another case of complementizer-clitic phonological interaction, where, in contrast

---

(i) *ħaaħaa qalam-ii /-i-ya*.  
 this pen(nom)-1sg.acc  
 ‘This is my pen.’

Note there is no case deletion with non-1sg possessive clitics; thus, the nominative surfaces in (ii) (Albuhayri 2013:71).

(ii) *ħaaħaa bayt-u-ki*  
 this house-nom-2sg.fem.acc  
 ‘This is your house.’

16 There are other complex complementizers that trigger cliticization which are also based on either *ʔinna* or *ʔanna*.



to the Standard Arabic cases discussed above where the form of the clitic is affected, the form of the complementizer is affected.

At any rate, what is important for us is that we have here a case of phonological interaction between a complementizer and the material in its complement. If such interaction is spell-out domain sensitive, i.e. if it cannot take place between elements in different spell-out domains, we have another argument here that what is sent to spell-out is not the complement of a phasal head.

In Bošković (2016c) I give another argument to this effect based on tone sandhi in Taiwanese, in particular, cases like (30), where tone sandhi, indicated by •, applies between the Comp *kong* and its IP complement. Tone sandhi, which Bošković (2016c) and Simpson and Wu (2002) show is spell-out domain sensitive, in (30) applies between a phasal head and its complement (see Bošković 2016c for an account of the full paradigm, including the data noted in Simpson and Wu 2002).

- (30) *A•hui liau•chun• kong• A•sin si• tai•pak• lang.*  
 Ahui thought C Asin is Taipei person  
 ‘A-hui thought that A-sin is from Taipei.’ (Simpson and Wu 2002:79)

To summarize, the phenomena discussed in this section illustrate the impossibility of PF interaction between the edge of phase XP and material outside of phase XP, as well as the possibility of PF interaction between the edge of phase XP and the complement of X, a state of affairs which is expected under phasal spell-out, but not under phasal complement spell-out.

It is also worth noting that early pre-phasal approaches to prosodic phrasing (e.g. Nespor and Vogel 1986; Selkirk 1986) have anticipated phasal spell-out, not phasal complement spell-out. Thus, the standard assumption in these approaches to prosodic structure, which is taken to be determined by (though it does not completely correspond to) the syntactic structure, is that the left edge of a CP corresponds to an intonational phrase boundary. In other words, the correspondence here is with a phase, not a phasal complement. This is natural if spell-out domains correspond to phases. If they don’t, we would have a rather strange situation where the prosodic domain would not correspond to what is sent to spell-out. What would be sent to spell out would be the IP below CP and the VP above the CP, but the correspondence would be with the “sandwiched” phrase, CP.

## 5. MULTIPLE SPELL-OUT AND SUCCESSIVE-CYCLIC MOVEMENT

I now turn to the interaction of multiple spell-out and successive cyclic movement (SCM), which raises an interesting problem. If both multiple spell-out and SCM were to be defined strictly on phases, phases would be spell-out units and SCM would target phases. A problem, however, would then arise. It is standardly assumed that what is sent to spell-out is no longer accessible to the syntax. Given this assumption, it is simply not possible to state the domain for both spell-out and SCM in terms of phases.



If SCM were to target spell-out units, the element undergoing movement would get frozen with the first step of SCM since it would be part of a spelled-out unit, hence no longer accessible for syntactic operations. Either spell-out or SCM can then be stated in terms of phases, but not both. What we in fact need is the following: In (31), where XP is the first phrase above YP and these phrases are affected by SCM and spell-out, YP should be the spell-out domain, and XP should be targeted by SCM (with YP spelled out after that movement). Furthermore, only one of these should correspond to phases.

(31) [<sub>XP</sub> [<sub>YP</sub>

The issue here is which of the two, XP or YP, should be a phase; i.e. the issue is whether the domain for SCM or the domain for spell-out should correspond to phases. In Chomsky (2001), XP is the phase; i.e. the domain for SCM is defined on phases, the domain for spell-out is not. In principle, it could be exactly the other way round, with YP being a phase, not XP. The domain for spell-out would then be defined on phases, instead of the SCM domain. In the next section, I will argue for this kind of a system; in particular, I will argue that this approach is preferable on both empirical and conceptual grounds to Chomsky's (2001) system, which is based on phasal complement spell-out. The main argument will involve a generalization concerning locality domains for movement established in Bošković (2015a). I will therefore first make a digression to discuss the generalization in question.

### 5.1 On the Complex NP Constraint and the lack of the Complex VP Constraint

Since Ross (1967) posited (32), illustrated by (33), where complex NP is a noun modified by a clause,<sup>17</sup> not much attention has been paid to an obvious contrast: While extraction from complex NPs is disallowed, extraction from such VPs is allowed, as shown by (34). In other words, there is Complex NP Constraint, but there is no Complex VP Constraint.

(32) The Complex NP Constraint (CNPC): Extraction from complex NPs is disallowed.

(33) \**How<sub>i</sub> did you hear* [<sub>NP</sub> *rumors* [<sub>CP</sub> *that* [<sub>IP</sub> *John bought a house t<sub>i</sub>]]]]?*

(34) *How<sub>i</sub> did you* [<sub>VP</sub> *think* [<sub>CP</sub> *that* [<sub>IP</sub> *a dog bit John t<sub>i</sub>]]]]?*

The standard research strategy within the Minimalist Program has been to focus on (34), putting (33) aside, with the theories of the locality of movement built on the basis of (34). Bošković (2015a), however, argues this research strategy has been misguided since (33) represents a pervasive pattern found in many contexts, while (34) is highly

---

17 I will ignore relative clauses, which involve extraction from adjuncts. (Note also that Safir 1985 shows that (32) cannot be reduced to the adjunct condition by treating nominal clausal complements as appositives/adjuncts.)

exceptional. Thus, extraction is banned not only from clausal, but all complements of Ns. Furthermore, APs, PPs, even ergative VPs pattern with NPs in this respect. I will briefly summarize the relevant data in the following section, restating the descriptive generalization reached in Bošković (2015a) based on these data in phase-theoretic terms. Section 5.2 will give a deduction of the generalization, where a phase is completely inaccessible to a higher phasal head, with no edge/PIC exception.<sup>18</sup> The crucial component of the deduction will be phasal spell-out, i.e. that what is sent to spell-out is phases, not phasal complements.

### 5.1.1 On the Complex XP Constraint<sup>19</sup>

The CNPC concerns clausal complements. However, extraction is banned not only from clausal complements of nouns, but all complements of nouns. Observe in this respect (36) and (38), which contrast with (35) and (37) (such contrasts were noted in Bach and Horn 1976 and Chomsky 1973).<sup>20</sup> Note that I assume a reanalysis/pruning account of P-stranding, where there is no PP in (37), hence like (35), (37) involves extraction of the nominal complement, in contrast to (38), which involves extraction out of it. (Section 5.4 actually gives an account of P-stranding where there is PP in (37) but its effects are voided so that for all intents and purposes *friends of* behaves like a complex head that takes  $t_i$  as its complement; pending section 5.4 I will put P-stranding aside.)

(35) *Of who(m)<sub>i</sub> did you see [friends t<sub>i</sub>]?*

(36) *??Of who(m)<sub>i</sub> did you see [<sub>NP</sub> enemies of friends t<sub>i</sub>]?*

(37) *Who<sub>i</sub> did you see [friends of t<sub>i</sub>]?*

(38) *?\*Who<sub>i</sub> did you see [<sub>NP</sub> enemies of friends of t<sub>i</sub>]?*

Consider now Greek, where both genitive DPs and PPs function as nominal complements. Both cases exhibit a simple/deep extraction contrast, extraction being disallowed from the nominal complement, as illustrated below for the former. (40), which involves extraction out of a nominal complement, contrasts with (39), which involves extraction of a nominal complement.

18 See Bošković (2015a) for an alternative account based on Chomsky's (2013) labelling system, and Bošković (2016a) for an account based on a contextual approach to phases where X functions as a phase only after it is embedded into structure. The accounts in question do not differ only theoretically, they are also not equivalent empirically (see Bošković 2016a); I will, however, not compare them here since that would go beyond the main goal of this article.

19 This section sums up some of the arguments from Bošković (2015a); see that work for additional arguments. Since weak islands are sometimes completely weakened with argument extraction, adjunct extraction is more reliable, hence will be used whenever possible (another interfering factor with argument but not adjunct extraction concerns reanalysis and phase collapsing from section 5.4, see Bošković 2015a). However, in English it can be tested only with clausal complements, even *\*From which city<sub>i</sub> did Peter meet [girls t<sub>i</sub>]* being disallowed (see Chomsky 1986).

20 Since we are dealing here with argument extraction the locality violations are weaker (note that only the relevant NP in the unacceptable examples is marked).

- (39) *tu vivliu<sub>i</sub> mu ipes pos dhiavases tin [kritiki t<sub>i</sub>]*  
 the.gen book.gen me said.2sg that read.2sg the review  
 ‘You told me you read the review of the book.’ (Horrocks & Stavrou 1987)
- (40) \**tu vivliu<sub>i</sub> mu ipes pos dhiavases (tin) [<sub>NP</sub> enstasi [<sub>DP</sub> tis kritiki t<sub>i</sub>]]*  
 the.gen book.gen me said.2sg that read.2sg the objection the.gen review.gen  
 ‘You told me you read the objection to the review of the book.’

Turning to French *combien*-extraction, while simple *combien*-extraction, where the DP is a verbal complement, is allowed, deep *combien*-extraction, where the DP is a complement of a noun, is not. The unacceptable (42) also involves extraction out of a nominal complement.

- (41) *Combien<sub>i</sub> a-t-il consulté [<sub>DP</sub> t<sub>i</sub> de livres]?*  
 how.many has-he consulted of books
- (42) \**Combien<sub>i</sub> a-t-il consulté [<sub>DP</sub> (plusieurs/des) [<sub>NP</sub> préfaces [<sub>DP</sub> t<sub>i</sub> de livres]]]*  
 how.many has-he consulted several/some prefaces of books  
 ‘How many books did he consult several/some prefaces of?’

Being rather liberal regarding extraction from NPs, Serbo-Croatian (SC) provides several relevant cases. SC allows extraction of adjectives (43) and NP-adjuncts (45). Importantly, both are blocked when the NP from which extraction occurs functions as a nominal complement, as in (44) and (46).

- (43) *Pametne<sub>i</sub> on cijeni [<sub>t<sub>i</sub></sub> prijatelje]*  
 smart he appreciates friends
- (44) \**Pametnih<sub>i</sub> on cijeni [<sub>NP</sub> prijatelje [<sub>t<sub>i</sub></sub> studenata]]*  
 smart he appreciates friends students  
 ‘He appreciates friends of smart students.’
- (45) *Iz kojeg grada<sub>i</sub> je Petar sreo [djevojke t<sub>i</sub>]?]*  
 from which city is Peter met girls
- (46) \**Iz kojeg grada<sub>i</sub> je Petar kupio [<sub>NP</sub> slike [djevojke t<sub>i</sub>]]?*  
 from which city is Peter bought pictures girl  
 ‘From which city did Peter buy pictures of a girl?’

A variety of extractions thus shows that extraction from the complement of a noun is disallowed. There is then nothing special about CPs here; extraction from a nominal complement is disallowed regardless of its category. We then have the Generalized Complex NP Constraint in (47).

- (47) Extraction out of nominal complements is disallowed.

Consider now other lexical heads, starting with adjectives and the CNPC context. Adjectives also display the effect in question: (48) is unacceptable if the adjunct modifies the embedded clause.

(48) \**How<sub>i</sub>/Why<sub>i</sub> are you* [<sub>AP</sub> *proud* [<sub>CP</sub> *that John hired Mary t<sub>i</sub>*]]?

Extraction is also banned from non-CP complements of adjectives, as the contrast in (49) shows.

- (49) a. *Of who(m)<sub>i</sub> is he* [*proud t<sub>i</sub>*]  
 b. \**Of who(m)<sub>i</sub> is he* [<sub>AP</sub> *proud of* [*friends t<sub>i</sub>*]]?

Returning to Greek, like nouns, adjectives can take either genitive DP or PP complements in Greek. In both cases, extraction from the complement of *responsible* is banned.

(50) \**Tu ktiriu<sub>i</sub>* [<sub>AP</sub> *ipefthinos* [*tu fotismu t<sub>i</sub> / gia to fotismo t<sub>i</sub>*]]  
 the.gen building.gen is.responsible the.gen lighting.gen/for the lighting  
 ‘the building he is responsible for the lighting of’

We then have the Generalized Complex AP constraint.

(51) Extraction out of adjectival complements is disallowed.

The same holds for PPs. (52) replicates the simple/deep extraction contrast from NPs/APs.

- (52) a. *Who<sub>i</sub> did you read* [*about t<sub>i</sub>*]  
 b. \**Of who(m)<sub>i</sub> did you read* [<sub>PP</sub> *about friends t<sub>i</sub>*]]?

Prepositions can take CP complements in Spanish. Significantly, they also disallow extraction.<sup>21</sup>

(53)  
 \**¿cómo<sub>i</sub> se acordó* [<sub>PP</sub> *de* [<sub>CP</sub> *que* [*Pedro preparaba la comida t<sub>i</sub>*]]]  
 how clitic (s)he.remembered of that Pedro prepared.imperf the food

Greek (54) confirms the existence of the Generalized Complex PP Constraint, given in (55).

(54) \**Tinos<sub>i</sub> endhiaferese* [<sub>PP</sub> *ya* [*ti fili t<sub>i</sub>*]]  
 who.gen be.interested.2sg for the friend  
 ‘Whose friend are you interested in?’ (Horrocks & Stavrou 1987)

(55) Extraction out of complements of prepositions is disallowed.

21 Some languages treat (some) Ps as inherent Case-markers (see Nunes 2009). Such Ps are not relevant to our concerns.

Before we address the obvious question, why are VPs different here, it is important to note that they are not always different. The effect in question is actually found with ergative VPs. Thus, (56), involving a non-ergative verb, is better than (57), involving an ergative verb.

- (56) *Who did they see (some) friends of  $t_i$  yesterday?*  
 (57) *?\*Who<sub>i</sub> did there arrive (some) friends of  $t_i$  last week?*

Turning to ergatives taking CP arguments, consider (58)-(59). They involve uncontroversially ergative psych verbs, where the CP is located in the V-complement position (see Belletti & Rizzi 1988; Pesetsky 1995; Landau 2010).<sup>22</sup> Importantly, both argument and adjunct extraction are degraded here, the latter being worse, as expected.<sup>23</sup>

- (58) a. *??What<sub>i</sub> did it appeal to Mary [that John fixed  $t_i$ ]?*  
       b. *\*How<sub>i</sub> did it appeal to Mary [that John fixed the car  $t_i$ ]?*  
 (59) a. *??What<sub>i</sub> did it depress Mary [that John sold  $t_i$ ]?*  
       b. *\*How<sub>i</sub> did it depress Mary [that John was fired  $t_i$ ]?*

There are also transitive ergatives that do not take CP arguments, where only argument extraction can be checked (see footnote 19). Extraction is also degraded in such cases (see also Belletti and Rizzi 1988 for Italian).

- (60) *??Who<sub>i</sub> did your behavior bother [the sister of  $t_i$ ]?* (Johnson 1992)  
 (61) *?\*Who<sub>i</sub> did John's embarrassment escape [friends of  $t_i$ ]?*

The Generalized Complex VP Constraint effects thus emerge with ergative verbs.

- (62) Extraction out of complements of ergative verbs is disallowed.

When properly generalized, the CNPC thus represents a pervasive pattern found in many contexts. Extraction is banned not only from clausal but all nominal complements. APs, PPs, and ergative VPs pattern with NPs.<sup>24</sup> In other words, with the exception of non-ergative Vs, extraction is banned from complements of lexical heads.

- (63) **The Complex XP Constraint** (where  $X \neq$  non-ergative V)  
 Extraction out of complements of lexical heads is disallowed.

22 For ergatives with just CP arguments, see Bošković (2015a), where it is shown that when some interfering factors are controlled for they also show Complex VP Constraint effects.

23 (58)-(59) may involve short V-movement, which may exist in English independently of  $v$  (Johnson 1991; Lasnik 1999)

24 See Bošković (2015a) on passives.

### 5.1.2 Restating the Complex XP Constraint

This section restates (63) within the phase theory of Bošković (2015a). A deduction of the phase-theoretic restatement of (63) will then be presented that is crucially based on phasal spell-out.

In Chomsky (2000), certain phrases are always phases regardless of their syntactic context. Many have, however, argued that the phase status of X can be affected by its syntactic context. Thus, Bošković (2013a, 2014) argues that the highest projection in the extended domain of a lexical head functions as a phase. There, vP is a phase as the highest projection in the extended domain of V. There is a phase even with ergatives even if vP, which is responsible for external  $\theta$ -role assignment, is absent; VP is then the phase as the highest projection in the domain of V. One issue, however, that arises with this approach concerns what counts as the extended domain of a lexical category. Consider V. Bošković (2013a) takes vP to close the extended domain of V. Wurmbrand (2014) proposes a modification of Bošković's (2013a) approach where AspectP is also included in this domain, which was also adopted in Bošković (2014). Grimshaw's (1990) approach to extended projections would extend the domain of V even further. What this shows is that determining exactly what counts as the extended domain of a lexical category (at least with verbs) is not straightforward. Bošković (2015a), however, proposes a modification that eschews this issue. Taking as the point of departure Grohmann (2003), where a clause is divided into three domains, the discourse, the agreement, and the  $\theta$ -domain, Bošković (2015a) proposes to collapse the agreement and the discourse domain into one domain, giving us two domains: thematic and non-thematic. What is relevant for our purposes here is the former. The proposal is that the thematic domain of each lexical category corresponds to a phasal domain, where the highest projection in the domain functions as a phase, as in Bošković (2013a). The departure from Bošković (2013a, 2014) here is that the thematic domain replaces the murkier notion of extended domain. vP is then a phase as the highest projection in the thematic domain. With ergatives, due to the lack of vP, VP is the highest projection in the thematic domain hence a phase. Notice that the presence of a non- $\theta$ -marking vP with ergatives would not affect anything here: VP would still be a phase (for this reason I will ignore the issue in question below). This is in contrast to Bošković (2013a, 2014), where the presence of a non- $\theta$ -marking vP would actually void VP of phasehood.<sup>25</sup>

Now, assuming the approach to phasehood where the highest projection in the thematic domain of every lexical head functions as a phase, it turns out that all the examples that instantiate (63) actually involve the context in (64), where a phasal head takes a phase as its complement.<sup>26</sup>

---

25 I assume that whether or not an external  $\theta$ -role is to be assigned is indicated in the  $\theta$ -grid of the verb (even when it is assigned in SpecvP; see here Sawada 2015), which means that the phasal status of VP (i.e. whether or not VP is a phase) can be determined locally (depending on whether all the  $\theta$ -roles in the  $\theta$ -grid of the verb are assigned). Following standard assumptions, I also assume that a  $\theta$ -marked externally merged Spec of XP is created before successive-cyclic movement can target the edge of XP. The reader should bear these points in mind during the discussion below.

26 Following standard assumptions, I assume that CPs and DPs are also phases.

(64) [<sub>XP=Phase</sub> [<sub>YP=Phase</sub>]]

To illustrate, NP is a phase in (33) as the highest projection in the nominal thematic domain. The same holds for AP in (48) and PP in (53), where AP and PP are the highest projections in the A/P thematic domains. Focusing on the nominal case, the noun takes CP, which is a phase, as its complement in (33). This means that (33) involves the double-phase context from (64). The same holds for Greek (40), where there is a DP phase right below the NP phase. In fact, all the cases of the Generalized Complex NP Constraint discussed above involve (64). The same holds for the Generalized Complex AP Constraint. The adjective, a phase head, takes a CP phase complement in (48) and a DP/PP phase complement in Greek (50). The same holds for the Generalized Complex PP Constraint. The preposition, a phasal head, also takes a phasal complement in all the relevant cases (see (52)-(54)). Consider finally the VP cases, i.e. (56)-(61). As discussed above, ergative verbs behave differently from non-ergative verbs in that they show Complex XP Constraint effects (i.e. there is a Generalized Complex VP effect with ergative verbs). The obvious conclusion is that vP is what matters here, which follows from the current phasal system. With non-ergative verbs, vP is the highest projection in the verbal thematic domain. This means VP is not a phase with non-ergative verbs. As a result, extraction from clausal complements of non-ergative verbs, as in (65), does not involve (64). In contrast, ergatives lack the thematic vP layer. This means that VP is the highest (and only) projection in the relevant thematic domain hence a phase in (66). (66) then involves a double-phase configuration. (Phases are given in boldface. For ease of exposition, I ignore V-movement here; recall also that the potential presence of a non- $\theta$ -marking vP would not affect anything here, hence it is ignored below).

(65) *How<sub>i</sub> did you* [<sub>VP</sub> *t<sub>i</sub>* [<sub>VP</sub> *think* [<sub>CP</sub> *t<sub>i</sub>* *that* [<sub>IP</sub> *John* [<sub>VP</sub> *fixed the car t<sub>i</sub>*]]]]]]]?

(66) \**How<sub>i</sub> did it* [<sub>VP</sub> *t<sub>i</sub>* *appeal to Mary* [<sub>CP</sub> *t<sub>i</sub>* *that* [<sub>IP</sub> *John* [<sub>VP</sub> *fixed the car t<sub>i</sub>*]]]]]]?

Extraction is thus disallowed in the configuration in (64), where a phasal head takes a phase as its complement. (63) can then be restated as in (67). ((67) will be slightly revised below).

(67) The Phase-over-Phase Constraint: Extraction is banned from phases that function as complements of phasal heads (i.e. the double-phase configuration in (64)).

Before proceeding with the account of (67), it is worth pausing here to address the obvious question the current discussion raises: why is there no Complex VP Constraint, in contrast to the Complex NP Constraint, the Complex AP Constraint and the Complex PP Constraint. A clue here is provided by the existence of the Complex VP Constraint effect with ergatives. The obvious difference between ergative and non-ergative verbs is the existence of ( $\theta$ -marking) vP with the latter. (67) capitalizes on this: the current approach to phases yields a principled difference (deduced below) between ergative



and non-ergative verbs given the presence of vP with the latter. Generalizing the VP situation, the reason for the contrast between acceptable (34) and ill-formed (33)/(48)/(53)/(58b), i.e. the reason for the different behavior of non-ergative VP and NP/AP/PP/ergative VP regarding the Complex XP Constraint, is the presence of vP, i.e. the assignment of the external  $\theta$ -role in a projection distinct from VP. There is then no such projection with NP/AP/PP.<sup>27</sup>

(67) restates (63) in phase-theoretic terms. The two are actually not equivalent empirically. Although all the cases discussed so far are captured by both (63) and (67), there are cases that can tease them apart. One such case involves AP extraction. Recall SC bans deep AP extraction, as in (44). Both (63) and (67) capture (44), the relevant structure of which prior to AP extraction is given in (68) (following Bošković 2012, I assume that SC lacks DP (though this is not crucial here)).

(68) ...<sub>[NP prijatelje</sub> [<sub>[NP pametnih studenata]]  
 friends smart students<sub>GEN</sub></sub>

The noun here assigns genitive, which is the structural case assigned by nouns in SC (Franks 1994; Bošković 2013a). However, nouns that assign inherent case allow deep AP extraction (deep adjunct extraction from (46) is also allowed with inherent case assigning nouns, see Bošković 2013a).

(69) ?*Kakvom<sub>i</sub> ga je uplašila [prijetnja [t<sub>i</sub> smrću]]?*  
 what-kind-of him is scared threat death<sub>INSTR</sub>  
 ‘Of what kind of death did a threat scare him?’ (Bošković 2013a)

Bošković (2013a) argues that inherent case-assignment involves a dummy linker-like projection FP, as in the structure in (70).<sup>28</sup> The higher noun here takes FP as its complement, with F in turn taking the lower NP as its complement.

(70) ...<sub>[NP prijetnja</sub> [<sub>FP</sub> [<sub>[NP kakvom smrću]]]]</sub>

Notice now that (69) still violates (63) since it involves extraction from the complement of a lexical head. However, assuming that due to its nature the dummy linker projection is not a phase, we don’t have a phase-over-phase configuration here. The case in question then differentiates (63) and (67), favoring (67) (but see Talić 2013; note also that (69) is consistent with the deduction of (67) below).

27 nP/pP/aP are often posited for the sake of uniformity with VP. But the fact is that there is no uniformity here at all regarding extraction. Notice, however, that n/p/aP can still exist, they just would not be part of the thematic domain (i.e. they would not be assigning a  $\theta$ -role; see Bošković 2015a for relevant discussion).

28 Bošković (2013a) basically follows here the long-standing intuition that inherent case assignment involves a preposition-like element, which is implemented through the presence of FP in (70).



Having restated (63) in phase-theoretic terms, which turned out to be preferable to (63) on empirical grounds (see footnote 39 for another relevant case), I turn to the deduction of (67).

## 5.2 Deducing the ban on extraction in phase-over-phase configurations

A clue for deducing (67) can be found in Chomsky's (2001) approach to the PIC. While in Chomsky (2000), when a phase is assembled only its edge is accessible to anything outside of the phase (due to the PIC), for Chomsky (2001) the PIC effect does not kick in immediately; it kicks in only when the next phase head is merged. As a result, the complement of a phase head is accessible to the next head if that head is not a phasal head. However, it is inaccessible if that head is a phasal head. This kind of distinction between phasal and non-phasal heads is exactly what we need. Recall that extraction from a phase is disallowed if the phase is merged with a phasal head, but not if it is merged with a non-phasal head. What we then need is to modify Chomsky's (2001) approach to the PIC in a way that still makes a difference between phasal and non-phasal heads regarding the accessibility of a phase they are merged with. In particular, the edge of a phase needs to be accessible only to non-phasal heads; a phase needs to be completely inaccessible to the next phasal head. This can be captured if the PIC holds only for non-phasal heads. It can then be used as a gate for movement only if the head merged with a phase is a non-phasal head. This means that the edge of phase XP in (71) is accessible to the non-phasal head Y, but not to the phasal head Z.

(71)  $Z(Y) [_{XP} \text{ wh } X]$

In (33)/(72), *how* then moves to the embedded SpecCP, CP being a phase. However, since CP is now completely inaccessible when the next phasal head is merged, when N, a phasal head, is merged, *how* can no longer move.

(72)  $N [_{CP} \text{ how } C [_{IP} \dots]]$

*How* also moves to the embedded SpecCP in (34). However, the head merged with the CP in (34)/(73a) is not a phasal head. This means that *how*, located at the CP phase edge, is accessible for movement to V. V then attracts *how*, which moves to SpecVP. Since VP is not a phase, in contrast to *how* in (72), *how* in (73b) is accessible to the next phasal head (*v*), hence can move to its edge.

(73) a.  $V [_{CP} \text{ how } C [_{IP} \dots]]$   
 b.  $v [_{VP} \text{ how } V [_{CP} \text{ how } C [_{IP} \dots]]]$

While the account works, it is rather stipulative: why would the PIC hold only for non-phasal heads?

Recall, however, what is needed here: a phase needs to be completely inaccessible at the next phasal level, i.e. when the next phasal head is merged, with no PIC loophole.

There is actually a rather natural way of accomplishing this if we assume that what is sent to spell-out is phases, not phasal complements. Furthermore, we will see that the phasal spell-out account enables us to dispense with the PIC completely (in contrast to the analysis outlined above).

What we need here is to close the PIC loophole in one context. The most natural way of closing that loophole is to eliminate the PIC loophole altogether. Suppose then that there is nothing like the PIC in the syntax. Under multiple spell-out, pieces of syntactic structure are transferred to the interfaces during the derivation. Assuming that what is transferred to the interfaces is no longer accessible in the syntax, there is no need to have anything like the PIC holding as a principle in the syntax itself, structure that is spelled-out will anyway be inaccessible in the syntax.

Recall now that in the original proposal regarding phases, what was sent to spell-out was the phase itself. The assumption was later modified in that what was sent to spell-out was the phasal complement. As discussed above, this resulted in a rather strange situation: while theoretically, phases are the crucial units in the multiple spell-out framework, for all practical purposes the crucial units are actually phasal complements. But, in contrast to phases, phasal complements have no theoretical status, as can be easily seen by comparing the great deal of effort that has gone into coming up with the proper definition of what counts as a phase with no attempts whatsoever of that kind regarding phasal complements. The reason is of course simple: phasal complements have no theoretical status, only phases do. They should then play no role in spell-out, what is transferred to spell-out should be phases. Recall also that arguments for multiple spell-out regarding PF phenomena favor phasal spell-out (see section 4), that phasal spell-out fits more naturally with the labeling theory, where labeling is interface-driven and what is being labeled is phases, and that it is also simpler in that only phases are ever sent to spell-out under the phasal spell-out approach while under the phasal complement spell-out approach spell-out still targets phases with matrix clauses.

Let us then assume that what is sent to spell-out is phases, not phasal complements. However, following Chomsky (2001) (but adapting it to phasal spell-out), a phase is transferred to spell-out when the next phasal head enters the structure. In particular, following Bošković (2014), the transfer takes place as soon as the next phase head is merged. This analysis, which does not require the PIC at all and privileges phases, not phasal complements, for spell-out, accounts for (67). The gist of it is that the *wh* in (74) is accessible to *Y*, a non-phasal head, but not to *Z*, a phasal head, because merger of *Z* triggers immediate spell-out of the *XP* phase.

(74)  $Z(Y) [_{XP} wh]$

Let us apply this analysis to concrete cases. (33), a CNPC case, is straightforward. As soon as  $\bar{N}$ , a phasal head, is merged, CP is sent to spell-out. As a result, nothing within CP is accessible for movement from CP, hence *how* cannot move out of it (it doesn't actually matter whether *how* moves to SpecCP or not).

(75) N [<sub>CP</sub> ...how...] (\*How<sub>i</sub> did you hear [<sub>NP</sub> rumors [<sub>CP</sub> that John bought a house t<sub>i</sub>]])

In (34), CP merges with V, which is not a phasal head, in contrast to the head CP merges with in (33). CP is then not sent to spell-out in (34)/(76a). This means that *how* is accessible for movement to V. (If its base position is above vP, *how* is accessible to V in its base position.) *How* then moves to SpecVP.<sup>29</sup> Merger of v triggers spell-out of the lower phase, CP. However, since *how* has already moved out of it, it is not affected by the CP spell-out in (76), in contrast to (75).

(76) a. V [<sub>CP</sub> ...how...] (How<sub>i</sub> did you [<sub>VP</sub> think [<sub>CP</sub> that a dog bit John t<sub>i</sub>]])  
 b. v [<sub>VP</sub> how V [<sub>CP</sub>

The contrast between (33) and (34) is thus captured. The analysis extends to all other cases of (67).

The account has interesting architectural consequences. What is sent to spell-out is full phases. However, what is targeted by successive-cyclic movement (SCM) is not phases, but phrases above them. (In (76), the movement targets the first phrase above vP (see Bošković and Lasnik 2003; Lahne 2008; Den Dikken 2009a for arguments for such movement) and VP; unless additional assumptions are adopted, movement need not pass through phasal edges.)<sup>30</sup>

29 See here Rackowski and Richards (2005) and den Dikken (2009a,b), who also argue that successive-cyclic movement in (34) does not go through SpecCP (for them, the movement also targets the VP domain above CP; note also that the standard reconstruction arguments for a phase in the VP domain with both unergative and ergative verbs of the kind discussed in Fox (1998) and Legate (2003) extend to the current approach to the locality-of-movement in this domain).

30 There are many empirical arguments for successive cyclic movement, which quite conclusively show that movement proceeds in this manner. While that much is clear, it is much harder to use the relevant tests to determine the exact landing sites. Thus, while the cases where movement has a morphological reflex on the verb (which is often the case) argue for successive cyclic movement through the VP domain, as is crucially the case under the current analysis, it is difficult to use them to pin-point its exact landing site. There are also cases which have been used to argue that movement proceeds via SpecCP. While nothing would go wrong in the current system if such movement occurs, the system does not require it. Importantly, all the relevant cases have been quite convincingly argued not to involve successive-cyclic movement through SpecCP; i.e. a closer scrutiny of such cases has revealed that they either involve terminal movement to SpecCP, no movement at all, or successive-cyclic movement via positions other than SpecCP (e.g. the well-known case of agreeing intermediate Cs in Kinande turns out not to involve movement; all movement diagnostics, like reconstruction and islandhood, fail with such cases; some cases that have been traditionally assumed to involve intermediate wh-agreement do not even exhibit such agreement; thus in many languages wh-movement affects the agreement relationship holding between the verb and the intermediate complementizer—it is not the case that the wh-phrase itself agrees with the C; at any rate, there are quite a few works arguing that languages that have been traditionally claimed to involve Spec-Head agreement between a wh-phrase and an intermediate C have been misanalyzed, see e.g. Boeckx (2004), Bošković (2008), Schneider-Zioga (2009), Rackowski and Richards (2005), Lahne (2008), Finer (2003), Noonan (1999) and especially den Dikken (2009a,b). Regarding other types of more direct arguments for successive cyclic move-

An appealing property of the phase system is that phases are relevant to many phenomena. However, as discussed above, there is a problem with spell-out and SCM in this respect. Given that what is sent to spell-out is no longer accessible to the syntax, it is simply not possible to state the domain for both spell-out and SCM in terms of phases. If SCM were to target spell-out units, the moving element would get trapped and prevented from further movement since it would be part of a spelled-out unit. Only one of the two mechanisms, spell-out or SCM, can then be stated in terms of phases: for a moving element not to get caught in a spell-out unit, either the domain for SCM or the domain for spell-out can correspond to phases, but not both. This in fact is the property of both Chomsky (2001) and the current system. They both have the following property: XP is sent to spell-out and movement targets YP right above it. In both systems, one of the two is defined on phases. The difference between the two is which mechanism is defined on phases. For Chomsky, it is SCM: SCM targets phases, spell-out doesn't.<sup>31</sup> In the current system, spell-out targets phases, SCM doesn't. In this respect, the systems seem equal conceptually. What we are, however, dealing with here is an issue of primacy: what should be privileged, spell-out or SCM. By defining the former on phases, with SCM piggy-backing on it, the current system privileges spell-out. Chomsky's system, on the other hand, privileges SCM. Many have, however, argued that SCM takes place so that the moving element escapes being sent to spell-out, see e.g. Bošković (2007), Fox & Pesetsky (2005), Stjepanović & Takahashi (2001); the intuition is in fact present even in Chomsky's analysis (furthermore, no one has ever argued that spell-out depends on SCM in this manner). This in itself argues for a system where spell-out is privileged, i.e. for a system like the one argued for here where spell-out is defined on phases, but SCM is not. Furthermore, the current system does not need anything like the PIC, which is needed in Chomsky's system. All we have is the assumption that phases are sent to spell-out, with SCM taking place so that the moving element avoids being sent to spell-out.<sup>32</sup>

---

ment via SpecCP, like Afrikaans (i), where the preposition was assumed to be stranded in the intermediate SpecCP during successive cyclic movement, the embedded verb being located in C (see Du Plessis 1977), they have also been shown not to provide evidence for such movement (see Den Dikken 2009a,b and references therein). Thus, the fact that in (ii), where *dink* does not move to C, the stranded preposition must precede, not follow, the verb shows that the preposition is not stranded in SpecCP, but in a position above it in the higher clause (which fits the current analysis); see den Dikken (2009a), Den Besten (2010), and Rackowski and Richards (2005) for additional evidence against the stranding-P-in-SpecCP analysis of (i).

(i) *waar/wat dink julle voor werk ons?*  
 where/what think you.PL for work we (Den Dikken 2009a:97)

(ii) *ek sou graag wou weet waar julle <voor> dink <\*voor> dat ons werk.*  
 I would gladly wanted know where you for think for that we work  
 'I would like to know what you think we work for.' (Den Dikken 2009a:100)

31 Recall, however, that under Chomsky's analysis spell-out does target phases with matrix clauses. Under the current analysis, spell-out is consistent in that it always targets only phases.

32 The analysis thus most naturally fits with systems where the driving force for successive-cyclic movement is implemented in this way, as in e.g. Bošković (2007).

At any rate, what is important for our purposes is that phase X is completely inaccessible when the next phasal head is merged. Movement from X is then possible only if X is first merged with a non-phasal head, which can “pull” the moving element out of X before the next phasal head enters the structure. *Wh* can then move out of CP in (77) only in the absence of Y.

(77)  $\mathbf{H} \text{ L } (\mathbf{Y}) [_{\text{CP}} \mathbf{C} [_{\text{IP}} \text{ wh}]]$

Extraction is thus banned in phase-over-phase configurations, which deduces (67).

An important point is in order. While extraction is disallowed from phases that function as phasal complements, such phases themselves can move (unless independent factors interfere, as with the movement of a CP complement of N). This also quite naturally falls out from the current analysis, which dispenses with the PIC. The conception of spell-out adopted here is essentially Uriagereka’s (1999) original conception of spell-out, which also does not have the PIC. Crucially, Uriagereka argues that when a phrase is sent to spell-out, nothing within it is available for further syntactic operations but the phrase itself is available. In his terms, sending A to spell-out, which results in establishing word order within A, turns A essentially into a compound/lexical item whose internal structure is inaccessible to the syntax. A itself is, however, accessible. As a result, while movement from the nominal complement in (40) is disallowed (we are dealing here with a phase that functions as a phasal complement), movement of the complement itself, as in (39), is allowed. When N merges with its complement, it triggers its spell-out. Nothing within the complement is then accessible for movement. However, the complement itself is accessible. Thus, in (78), the structure of (39)-(40) (I adopt the standard assumption that there is more than one phrase in the NP-functional domain), merger of N with K triggers spell-out of K. K is still accessible to X, a non-phasal head (though nothing within K is), hence K can move to SpecXP. Merger of D triggers spell-out of NP; however, this does not matter since K has already moved outside of NP.

(78)  $[_{\text{DP}} [_{\text{XP}} \text{ X } [_{\text{NP}} \text{ N } \text{ K}]]]$

### 5.3 Stress assignment in English: the Bresnan/Legate argument for multiple spell-out

Now that the details of the current approach to successive-cyclic movement and spell-out, where spell-out targets phases and successive-cyclic movement targets non-phasal projections above phases, have been laid down I return to stress assignment. In particular, we will see in this section that Legate’s (2003) discussion of primary stress assignment in English provides evidence for the current approach to successive-cyclic movement and spell-out.

Consider (79)-(81), where the relevant word bearing primary stress is underlined.

- (79) a. *Mary fixed the bike.*  
 b. *Mary fixed it.*

(Legate 2003:511)

- (80) a. *Mary liked the proposal that George leave.*  
 b. *Mary liked the proposal that George left.* (Bresnan 1972:75)
- (81) a. *Please put away the dishes.*  
 b. *?Please put the dishes away.* (Legate 2003:512)

Since the discussion in this section is based on a generalization regarding primary stress assignment reached in Legate (2003), I will couch it within the approach to stress assignment in English adopted by Legate. Legate assumes the Nuclear Stress Rule (NSR), where the primary stress in English is assigned to the final stress bearing element in the VP (i.e. the rightmost element), hence the contrast in (79).<sup>33</sup> The NSR is also responsible for stress assignment in (80a). As for (80b), the example represents one of the original arguments for multiple spell-out. In particular, based on such examples, Bresnan (1972) argues the NSR applies cyclically: in (80b), it applies before *proposal* moves from the most embedded object position (see Vergnaud 1974; Kayne 1994), assigning stress to *proposal* (in the current theory, this would be handled in terms of multiple spell-out). In (81a), the NSR applies normally, assigning stress to *dishes*. Importantly, in (81b), which is generally assumed to involve movement of *the dishes*, in particular, object shift (see Johnson 1991; Lasnik 1999, 2001; Gallego and Uriagereka 2007), the NSR can assign stress to *away*. Apparently, in (80b) the NSR applies before the object undergoes movement, assigning stress to it, but in (81b) the NSR applies after the object undergoes movement, hence it does not assign stress to it. The question is how this difference regarding stress assignment in (80b) and (81b) can be captured. As noted above, *the dishes* in (81b) is generally assumed to undergo object shift, moving to SpecvP. Under the standard analysis, where vP is a phase and successive-cyclic movement targets phasal edges, the object in (80b) also moves to SpecvP. After this movement, the VP complement of v is sent to spell-out in both (80b) and (81b). (80b) and (81b) then have the same derivation in all relevant respects (see (82), which gives the relevant part of the structure (recall  $t_i$  is sentence final, hence *the dishes/proposal* is in the position to be assigned stress by the NSR prior to the movement)). It is difficult to differentiate (80b) and (81b) with respect to stress assignment under this analysis, which relies on the standard assumptions regarding phases, successive-cyclic movement, and spell-out.

- (82) [<sub>vP</sub> the dishes/proposal<sub>i</sub> [<sub>vP</sub> ...t<sub>i</sub>]]

33 Kratzer and Selkirk (2007) do not appeal to the NSR in their analysis of German (see section 4). There is a good deal of discussion in the literature regarding whether the NSR applies in German, and if so, in which way (to capture the relevant differences between German and English), see e.g. Zubizarreta (1998) (there are conflicting data claims in Kratzer and Selkirk 2007 and Zubizarreta 1998 however). Since determining the precise implementation of the stress assigning mechanism (and capturing crosslinguistic variation in this domain) is beyond the scope of this article, I simply use the mechanisms the works cited here rely on (for a multiple spell-out approach to the NSR, see also Adger 2007).



How can the contrast between (80b) and (81b) then be accounted for? Interestingly, Legate (2003) proposes an account which is inconsistent with the standard assumptions regarding successive-cyclic movement and spell-out, discussed above regarding (82), but fits perfectly with the current approach. Legate argues that the crucial difference between (80b) and (81b) is that object movement in (80b) takes place to a position outside of the lowest spell-out domain, while in (81b) the object moves to a position within the lowest spell-out domain; in other words, she argues that there is only one copy of the object in the input to PF in the first phase of (80b), while there are two copies of the object in the input to PF in the first phase of (81b). As a result, when the PF operation that deletes non-initial copies within a spell-out domain applies in (81b), it deletes the lower copy of the object. The NSR then assigns stress to *away*, the rightmost element within the spell-out unit. Turning to (80b), since there is only one copy of the object in the first spell-out domain of this construction, the PF deletion operation that deletes non-initial copies, which treats each phase as a separate unit, does not delete this element when it applies to the first spell-out domain. The NSR then applies, assigning stress to *proposal*, as the rightmost element within this spell-out domain.<sup>34</sup>

At any rate, putting the details of the implementation aside, what's important here is Legate's proposal that what differentiates (80b) and (81b) is that the object in (80b) moves outside of the lowest spell-out domain, while in (81b) it moves within it, as a result of which there is only one copy of the object in the first spell-out domain of (80b), while there are two such copies in the first spell-out domain of (81b). This is in fact exactly what happens on the current analysis. As discussed above, *the dishes* undergoes object shift to SpecvP in (81b), with vP being a phase. However, object movement in (80b) does not target the same position. In contrast to the standard analysis, where successive-cyclic movement targets phasal edges, in the current analysis successive-cyclic movement targets phrases above phases. The object in (80b) then does not move to SpecvP, but to the Spec of the first phrase above vP.<sup>35</sup> This gives us (83)-(84) for (80b) and (81b) respectively.

- (83) [<sub>XP</sub> the proposal<sub>i</sub> [<sub>vP</sub> [<sub>VP</sub> ...t<sub>i</sub>]]]  
 (84) [<sub>vP</sub> the dishes<sub>i</sub> [<sub>VP</sub> ...t<sub>i</sub>]]

(80b) and (81b) thus have very different derivations when it comes to object movement in the current analysis. Most importantly, the derivations of (80b) and (81b) in

34 At a later phasal/spell-out domain, this occurrence of *proposal* is deleted in favor of a higher occurrence, with the primary stress realized on this higher occurrence (i.e. the occurrence that is not deleted; see Legate 2003 for details).

35 Since the identity of this phrase is not important for our purposes, I simply use XP. Note that analyses that assume that at least in some cases English has object shift also assume that V in English moves, though not as high as in Romance. The implicit assumption here is that there is more than one phrase above vP in the inflectional domain; there is in fact a great deal of evidence that simple TP-over-vP structure is inadequate, see e.g. Belletti (1990), Stjepanović (1998), Cinque (1999), Bošković (2001) (regarding V-movement), Bobaljik and Jonas (1996) (regarding subject positions) and Bošković (2004) (regarding floating quantifiers).

the current system fit perfectly Legate’s account of the contrast between (80b) and (81b) regarding stress assignment. Recall Legate crucially argues that the object in (80b) moves to a position outside of the first spell-out domain, while the object in (81b) moves to a position within the first spell-out domain, as a result of which there is only one copy of the object in the first spell-out domain of (80b) while there are two such copies in the first spell-out domain of (81b). This is exactly what happens under the current analysis. Given that spell-out targets phases, what is sent to spell-out in both (83) and (84) is vP. The object then moves outside of the first spell-out domain in (80b)/(83) but not in (81b)/(84). The stress assignment facts discussed by Legate (2003) can then be interpreted as providing additional evidence for the approach to spell-out and successive-cyclic movement argued for here.

#### 5.4 CED effects

Before concluding, I will briefly discuss some consequences of the analysis from section 5.2., which crucially relies on phasal spell-out, for the locality of movement.

In addition to deducing the Phase-over-Phase constraint from (67), turning it into a theorem, the analysis also captures the traditional ban on extraction from subjects in SpecIP. Consider (85).

$$(85) \text{ } [_{CP}[_{IP} \text{ Subject } [_{I'} ]]]$$

Recall that a phase is completely inaccessible to the higher phase head. Since subjects are phases (possibly only DPs), the subject is inaccessible to C, the next phasal head, hence C cannot attract anything out of it. (85) does not literally involve a phase-over-phase configuration. IP, a non-phase, dominates the subject but not CP. However, this is not enough. What is needed is for a non-phasal head to intervene between the two phases, i.e. for a non-phasal head to c-command the lower phase so that it can probe into it, attracting the moving element out of it. Since I does not c-command the subject, it cannot attract anything out of it. Extraction from subjects in SpecIP is then banned.<sup>36</sup>

The analysis extends to the ban on extraction from adjuncts if adjuncts are adjoined to complements of phasal heads (VP and IP). Consider (86), where K is also a phase (CP, DP, or PP).

$$(86) \text{ } [_{VP}] \text{ } K \text{ } [_{VP}] \text{ } [_{VP}]$$

The only element that intervenes between the vP phase (which sends K to spell-out) and K is a VP-segment, which is irrelevant (V cannot attract anything out of K since V does not c-command K). Extraction from K is then impossible. The analysis

---

36 Subjects in SpecvP allow extraction (Stepanov 2007). Here, at least one non-phasal head intervenes between vP and CP, which can pull out a moving element from the subject before C is merged (note that since only merger of a phasal head (not a projection of a phasal head) triggers spell-out, subject in SpecvP is not sent to spell-out until C is merged).



thus captures the ban on extraction from adjuncts.<sup>37</sup> In fact, the ban on extraction from subjects and the ban on extraction from adjuncts are unified with the Complex XP Constraint, all three being reduced to the ban on extraction in phase-over-phase contexts. (67) should, however, be slightly revised in light of the above discussion.

- (87) The Phase-over-Phase Theorem: Extraction is banned from a phase that is immediately c-commanded by a phasal head (where X is immediately c-commanded by head Y if there is no head Z such that Z c-commands X but not Y).

## 5.5 Phase collapsing

Consider now why P-stranding does not matter in the contrast in (37)/(35) vs (38)/(36), i.e. why (37) (repeated as (88)) does not instantiate (47). Many have implemented the reanalysis approach to P-stranding in terms of P-incorporation (overt or covert). (37)/(88) then involves P-to-N movement.

- (88) *Who<sub>i</sub> did you see* [<sub>DP</sub> [<sub>XP</sub> *t<sub>i</sub>* [<sub>NP</sub> *friends of<sub>j</sub>* [<sub>PP</sub> *t<sub>j</sub>* *t<sub>i</sub>*]]]]?

Based on a number of cases, some of which are discussed below, Bošković (2015a) proposes phase collapsing for phasal projections headed by two phase heads (due to the movement of the lower phase head to the higher phase head): the two phases are collapsed into one, the lower phrase not being a phase.<sup>38</sup> PP is then not sent to spell-out in (88) (note that there is a feature on the P and N that drives the movement in question which also indicates that the phasehood of PP will be voided hence PP is not spelled-out when N is merged). As a result, when X enters the structure *who* can move out of NP. XP not being a phase, *who* is available for movement after D is merged.<sup>39</sup>

Note that (89) is still ruled out: merger of *of'* causes spell-out of DP; *who* is then stuck within it.

- (89) \**Who<sub>i</sub> did you see enemies* [<sub>PP</sub> *of'* [<sub>DP</sub> [<sub>XP</sub> *t<sub>i</sub>* [<sub>NP</sub> *friends of''<sub>j</sub>* [<sub>PP</sub> *t<sub>j</sub>* *t<sub>i</sub>*]]]]]]?

*Of'* behaves like a phase head at the point of merger, sending DP to spell-out. The reason for this may be that the noun that can void its phasehood (*enemies*) has not yet

37 Adjuncts in ergative constructions require additional assumptions which I cannot go into here due to space limitations.

38 While there is similarity between phase collapsing and phase sliding/extension (den Dikken 2007; Gallego and Uriagereka 2007), where head movement extends the phase to the next projection, the former is much more constrained—it arises only when a phasal head moves to a phasal head, hence it also cannot turn a non-phase into a phase (see Bošković 2015a for a comparison; Bošković 2015a also assumes that the moved phase head must be a sister to a segment of the higher phase head in the phase-collapsing configuration).

39 Note also that (88) is another case that teases apart (67) and (63)—it conforms with (67) but not (63). ((88) involves extraction out of the complement of a lexical head but not a phase-over-phase configuration, PP not being a phase.)

entered the structure (this doesn't affect (88)) or that P-incorporation in English occurs only under P-stranding, hence not for *of* in (89). There is reason to favor the latter analysis, where the phasehood of PP' is never voided.

Setswana, a Bantu language where the noun precedes all other DP elements, which is analyzed in terms of N-to-D movement (see Carstens 2010), does not display CNPC effects.

- (90) *Ke m-ang yo o utlw-ile-ng ma-gatwe a gore nšša e lom-ile?*  
 it C1-who C1Rel 2sgSM hear-perf-Rel C6-rumor C6SM that C9-dog C9SM bite-perf  
 'Who did you hear rumors that a dog bit?' (Bošković 2015a:651)

This follows from phase collapsing. Due to N-to-D, the matrix object is headed by two phase heads, D and N. (I assume XP is either not present in Setswana or it is present, with X moving to D and N to the X+D head.) Due to phase collapsing, NP is then not a phase. Crucially, N doesn't cause spell-out of CP, hence extraction from CP is possible. This means that already at the point of merger, N is treated as a non-phasal head. (The presence of a D feature on N, which drives N-to-D, is sufficient to determine locally that the phasehood of N will be voided.) The implication of this for (89) is that *of* in (89) does not incorporate; P-incorporation in English occurs only in P-stranding contexts, hence *of* is a phasal head in (89) (the alternative account of (89) noted above cannot capture (90)). Returning to (90), since the first phase above CP is DP, the CP is not sent to spell-out until D is merged. Since there is at least one non-phasal projection between CP and DP (NP and possibly XP), 'who' can move out of the CP phase before D enters the structure, triggering spell-out of the CP.

Recall now the deduction of the ban on extraction from adjuncts. The first head above KP in (91) is *v*, which is also a phasal head hence cannot attract anything from KP (since it triggers immediate spell-out of KP). The account makes a prediction. If *K* moves to *v*, KP phasehood will be voided due to phase collapsing. KP will then not be spelled-out until *C* is merged (triggering spell-out for the whole *vP* phase). Movement from KP should then be possible since there are several heads lower than *C* that can pull the moving element out of KP before *C* is merged.

- (91) [<sub>VP</sub>] KP [<sub>VP</sub>]<sub>VP</sub>]

The surprising prediction is borne out. Galician has a phenomenon of D-incorporation, which voids islandhood, as discussed in Uriagereka (1988) and Bošković (2013b). Extraction from adjuncts is banned in Galician, as in (92). However, the ban is voided with D-incorporation, exactly as expected under the current analysis (I assume D incorporates into *V+v*).<sup>40</sup>

40 D-incorporation does not rescue CNPC violations in Galician (see Bošković 2015a). This is expected: what is responsible for the CNPC effect is the phasehood of NP, which is not affected by D-incorporation.

- (92) *??de que semana<sub>j</sub> traballastedes* [<sub>DP</sub> *o* *Luns* *t<sub>j</sub>*]  
of which week worked the Monday  
‘Of which week did you guys work the Monday?’
- (93) *de que semana<sub>j</sub> traballastede-lo<sub>i</sub>* [<sub>DP</sub>[<sub>D'</sub> *t<sub>i</sub>* *Luns* *t<sub>j</sub>*]]

## 6. CONCLUSION

The article has argued that what is sent to spell-out is phases, not phasal complements. Phasal spell-out has obvious theoretical advantages over phasal complement spell-out. In contrast to phases, phasal complements have no theoretical status. Furthermore, the phasal complement spell-out approach still requires phasal spell-out in one context (matrix clauses), while on the phasal spell-out approach only phases are ever sent to spell-out. We have also seen that the PIC can be eliminated under phasal spell-out. Phasal spell-out also fits more naturally with the recent labeling system, where labeling is interface-driven and what is being labeled is phases, not phasal complements.

Several empirical arguments have also been presented in favor of phasal spell-out regarding syntax-phonology interaction, the relevant phenomena being cliticization in Bulgarian and Arabic, stress assignment in German and English, raddoppiamento fonosintattico in Abruzzese, and tone sandhi in Taiwanese. These phenomena were shown to illustrate the impossibility of PF interaction between the edge of phase XP and material outside of phase XP, as well as the possibility of PF interaction between the edge of phase XP and the complement of X, both of which are expected under phasal spell-out, but not under phasal complement spell-out. The standard assumption that CP corresponds to an intonational phrase boundary also fits more naturally with phasal spell-out.

Another argument concerned the Phase-over-Phase constraint, a broad generalization that covers a number of islands. Taking as the point of departure the Complex NP Constraint, we have seen that extraction is banned not only from clausal but all complements of nouns. Adjectives, prepositions, and ergative verbs pattern with nouns in this respect, extraction from the complement of a lexical head being possible only with non-ergative verbs. Adopting an approach to phases where the highest projection in the thematic domain of a lexical head functions as a phase, the ban on extraction from complements of lexical heads was recast as the ban on extraction in double-phase configurations, i.e. the ban on extraction from phases that are immediately c-commanded by a phasal head—the Phase-over-Phase Constraint. The constraint was also extended to the CED, i.e. the ban on extraction from subjects and adjuncts, unifying all of these under the ban on extraction in phase-over-phase configurations. A deduction of the Phase-over-Phase constraint was given where phase XP is completely inaccessible, with no edge/PIC exception, once a higher phase head enters the structure. The crucial component of the analysis was phasal spell-out, i.e. that what is sent to spell-out is phases, not phasal complements. As a result, nothing within phase XP is accessible to higher phase YP, given that merger of Y leads to immediate spell-out of XP. Since phase XP is completely inaccessible when the next phasal head is merged, movement

from it is possible only if XP is first merged with a non-phasal head, which can pull the moving element out of XP before the next phasal head is merged. In Complex XP Constraint and *CED* contexts, the first merged head is a phasal head – we are dealing here with phase-over-phase configurations, which disallow extraction.

An interesting consequence of this system is that while what is sent to spell-out is full phases, what is targeted by successive-cyclic movement is not phases, but phrases above phases. Phases are relevant to many phenomena. There is, however, a problem with spell-out and successive-cyclic movement in this respect. Assuming that what is sent to spell-out is no longer accessible to the syntax, it is not possible to state the domain for both spell-out and successive-cyclic movement in terms of phases. If successive-cyclic movement were to target spell-out units, the moving element would also be sent to spell-out hence would be frozen for any further movement. Only one of the mechanisms in question can then be stated in terms of phases. We have here the following situation: YP is sent to spell-out and movement targets XP right above it, and either YP or XP is a phase (with movement to XP taking place after YP is spelled out). Chomsky (2001) takes XP to be the phase while the current system takes YP to be the phase. As noted above, in addition to deducing the Phase-over-Phase Constraint, the current system allows us to dispense with the PIC. All we have is the assumption that phases are sent to spell-out, with successive-cyclic movement taking place so that the moving element avoids being sent to spell-out. To the extent that it is successful the analysis of successive-cyclic movement presented here thus provides another argument for phasal spell-out, a different kind of argument from those involving syntax-phonology interaction. The argument was also confirmed by certain facts concerning stress assignment.

## References

- ABELS, Klaus (2001) “The Predicate Cleft Construction in Russian.” In: S. Franks/T. Holloway King/M. Yadroff (eds), *Formal Approaches to Slavic Linguistics: The Indiana Meeting 2000*. Ann Arbor: Michigan Slavic Publications, 1–19.
- ABELS, Klaus (2003) *Successive cyclicity, anti-locality, and adposition stranding*. Doctoral dissertation, University of Connecticut.
- ADGER, David (2007) “Stress and phasal syntax.” *Linguistic Analysis* 33, 238–266.
- AGBAYANI, Brian (2000) “Wh-subjects in English and the vacuous movement hypothesis.” *Linguistic Inquiry* 31, 703–713.
- ALBUHAYRI, Salem. 2013. *The pronominal system in Standard Arabic: Strong, clitic and affixal pronouns*. MA Thesis. Arizona State University.
- AN, Duk-Ho (2007) *Syntax at the PF interface*. Doctoral dissertation. University of Connecticut.
- BACH, Emmon/George HORN (1976) “Remarks on ‘Conditions on Transformations’.” *Linguistic Inquiry* 7, 265–299.
- BENMAMOUN, Elabbas (2000) *The feature structure of functional categories: A comparative study of Arabic Dialects*. Oxford: Oxford University Press.

- BESTEN, Hans den (2010) “Is there ‘preposition stranding in Comp’ in Afrikaans? No way!” In: C. J.-W. Zwart/M. de Vries (eds), *Structure preserved: Studies in syntax for Jan Koster*. Amsterdam: John Benjamins, 57–64.
- BELLETTI, Adriana (1990) *Generalized verb movement*. Turin: Rosenberg and Sellier.
- BELLETTI, Adriana/Luigi RIZZI (1988) “Psych-verbs and theta theory.” *Natural Language and Linguistic Theory* 6, 291–352.
- BIBERAUER, Theresa/Roberta D’ALESSANDRO (2006) “Syntactic doubling and the encoding of voice in Eastern Abruzzese.” In: D. Baumer/D. Montero/M. Scanlon (eds), *Proceedings of WCCFL 25*. Somerville, M: Cascadilla Press, 87–95.
- BOBALJIK, Jonathan (2002) “A-Chains at the PF-interface: Copies and ‘covert’ movement.” *Natural Language and Linguistic Theory* 20.2, 197–267.
- BOBALJIK, Jonathan/Dianne JONAS (1996) “Subject positions and the roles of TP.” *Linguistic Inquiry* 27, 195–236.
- BOECKX, Cedric (2003) *Islands and Chains*. Amsterdam: John Benjamins.
- BOECKX, Cedric (2004) *Bare syntax*. Ms., Harvard University.
- BOŠKOVIĆ, Željko (1994) “D-structure, theta-criterion, and movement into theta-positions.” *Linguistic Analysis* 24, 247–286.
- BOŠKOVIĆ, Željko (2001) *On the nature of the syntax-phonology interface: Cliticization and related phenomena*. Amsterdam: Elsevier Science.
- BOŠKOVIĆ, Željko (2002a) “On multiple wh-fronting.” *Linguistic Inquiry* 33, 351–383.
- BOŠKOVIĆ, Željko (2002b) “Clitics as non-branching elements and the Linear Correspondence Axiom.” *Linguistic Inquiry* 33, 329–340.
- BOŠKOVIĆ, Željko (2004) “Be careful where you float your quantifiers.” *Natural Language and Linguistic Theory* 22, 681–742.
- BOŠKOVIĆ, Željko (2007) “On the locality and motivation of Move and Agree: An even more minimal theory.” *Linguistic Inquiry* 38, 589–644.
- BOŠKOVIĆ, Željko (2008) “On successive cyclic movement and the freezing effect of feature checking.” In: J. M. Hartmann/V. Hegedüs/H. van Riemsdijk (eds), *Sounds of silence: Empty elements in syntax and phonology*. Amsterdam: Elsevier North Holland, 195–233.
- BOŠKOVIĆ, Željko (2012) “On NPs and clauses.” In: G. Grewendorf/T. E. Zimmermann (eds), *Discourse and grammar: From sentence types to lexical categories*. Berlin: de Gruyter, 179–242.
- BOŠKOVIĆ, Željko (2013a) “Phases beyond clauses.” In: L. Schürcks/A. Giannakidou/U. Etxeberria (eds), *The nominal structure in Slavic and beyond*. Berlin: de Gruyter, 75–128.
- BOŠKOVIĆ, Željko (2013b) “Traces do not head islands: What can PF deletion rescue?” In: Y. Miyamoto et al (eds), *Deep Insights, Broad Perspectives: Essays in Honor of Mamoru Saito*. Tokyo: Kaitakusha, 56–93.
- BOŠKOVIĆ, Željko (2014) “Now I’m a phase, now I’m not a phase.” *Linguistic Inquiry* 45, 27–89.

- BOŠKOVIĆ, Željko (2015a) “From the Complex NP Constraint to everything: On deep extractions across categories.” *The Linguistic Review* 32, 603–669.
- BOŠKOVIĆ, Željko (2015b) “Deducing the Generalized XP Constraint from phasal spell-out.” In: M. Ziková/P. Caha/M. Dočekal (eds), *Slavic Languages in the Perspective of Formal Grammar: Proceedings of FDSL 10.5, Brno 2014*. Bern: Peter Lang, 79–100.
- BOŠKOVIĆ, Željko (2016a) “Contextual phasehood and the ban on extraction from complements of lexical heads: When does X become a phase?” In: M. Yasui/M. Mizuguchi (eds), *Phase theory and its consequences*. Tokyo: Kaitakusha, 5–39.
- BOŠKOVIĆ, Željko (2016b) “On the timing of labeling: Deducing Comp-trace effects, the Subject Condition, the Adjunct Condition, and tucking in from labeling.” *The Linguistic Review* 33, 17–66.
- BOŠKOVIĆ, Željko (2016c) *Tone sandhi in Taiwanese and phasal spell-out*. Ms., University of Connecticut.
- BOŠKOVIĆ, Željko/Howard LASNIK (2003) “On the distribution of null complementizers.” *Linguistic Inquiry* 34, 527–546.
- BOŠKOVIĆ, Željko/Jairo NUNES (2007) “The copy theory of movement: A view from PF.” In: N. Corver/J. Nunes (eds), *The copy theory of movement*. Amsterdam: Benjamins, 13–74.
- BRESNAN, Joan (1972) *Theory of complementation in English syntax*. Doctoral dissertation, MIT.
- CARSTENS, Vicki (2010) *Head-movement in Bantu DPs*. Presented at NELS 41, University of Pennsylvania.
- CHENG, Lisa/Laura DOWNING (2016) “Phasal Syntax=Cyclic Phonology?” *Syntax* 19, 156–191.
- CHOMSKY, Noam (1973) “Conditions on transformations.” In: S. Anderson/P. Kiparsky (eds), *A festschrift for Morris Halle*. New York: Holt, Rinehart and Winston, 232–286.
- CHOMSKY, Noam (1986) *Barriers*. Cambridge: MIT Press.
- CHOMSKY, Noam (1993) “A minimalist program for linguistic theory.” In: K. Hale/S. J. Keyser (eds), *The view from Building 20*. Cambridge: MIT Press, 1–52.
- CHOMSKY, Noam (2000) “Minimalist inquiries.” In: R. Martin/D. Michaels/J. Uriagereka (eds), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. Cambridge: MIT Press, 89–155.
- CHOMSKY, Noam. 2001. “Derivation by phase.” In: M. Kenstowicz (ed.), *Ken Hale: A life in language*. Cambridge: MIT Press, 1–52.
- CHOMSKY, Noam (2013) “Problems of projection.” *Lingua* 130, 33–49.
- CHOMSKY, Noam (2015) “Problems of projection: Extensions.” In: E. Di Domenico/C. Hamann/S. Matteini (eds), *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*. Amsterdam: John Benjamins, 3–16.
- COLLINS, Chris (2002) “Eliminating labels.” In: S. D. Epstein/T. D. Seely (eds), *Derivation and explanation in minimalist syntax*. New York: Wiley, 42–64.



- CINQUE, Guglielmo (1999) *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.
- D'ALESSANDRO, Roberta/Tobias SCHEER (2015) "Modular PIC." *Linguistic Inquiry* 46, 593–624.
- D'ALESSANDRO, Roberta/Ian ROBERTS (2010) "Past participle agreement in Abruzzese: Split auxiliary selection and the null-subject parameter." *Natural Language and Linguistic Theory* 28, 41–72.
- DIKKEN, Marcel den (2007) "Phase extension: Contours of a theory of the role of head movement in phrasal extraction." *Theoretical Linguistics* 33, 1–41.
- DIKKEN, Marcel den (2009a) "Arguments for successive-cyclic movement through SpecCP." *Linguistic Variation Yearbook* 9, 89–126.
- DIKKEN, Marcel den (2009b) "On the nature and distribution of successive cyclicity. Adjunction, resumption, and scope marking as the roads to success in long-distance relation building." Ms., CUNY Graduate Center, New York.
- DIESING, Molly (1992) *Indefinites*. Cambridge, MA: MIT Press.
- DIESING, Molly (1996) "Semantic variables and object shift." In: H. Thráinsson/S. Epstein/S. Peter (eds), *Studies in comparative Germanic syntax*. Dordrecht: Kluwer. 66–84.
- EPSTEIN, Samuel David (1999) "Un-principled syntax: The derivation of syntactic relations." In: S. Epstein/N. Hornstein (eds), *Working minimalism*. Cambridge: MIT Press. 317–345.
- EPSTEIN, Samuel David/Erich GROAT/Ruriko KAWASHIMA/Hisatsugu KITAHARA (1998) *A derivational approach to syntactic relations*. New York: Oxford University Press.
- FASSI FEHRI, Abdelkader (1988) "Agreement in Arabic, binding and coherence." In: M. Barlow/C. Ferguson (eds), *Agreement in natural language*. Stanford: CSLI, 107–158.
- FASSI FEHRI, Abdelkader (1993) *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer.
- FOX, Danny (1998) *Economy and semantic interpretation*. Doctoral dissertation, MIT.
- FOX, Danny/David PESETSKY (2005) "Cyclic linearization of syntactic structure." *Theoretical Linguistics* 31, 1–46.
- FINER, Daniel (2003) "Phases and movement in Selayarese." In: A. Rackowski/N. Richards (eds), *Proceedings of AFLA 8*. Department of Linguistics and Philosophy, MIT, 157–69.
- FRANKS, Steven (1994) "Parametric properties of numeral phrases in Slavic." *Natural Language and Linguistic Theory* 12, 570–649.
- FRANKS, Steven (1998) *Clitics in Slavic*. Paper presented at the Comparative Slavic Morphosyntax Workshop, Spencer Creek, Indiana, June 1998.
- FRANKS, Steven/Željko BOŠKOVIĆ (2001) "An argument for multiple spell-out." *Linguistic Inquiry* 32, 174–183.

- FRANKS, Steven/Tracy KING (2000) *A handbook of Slavic clitics*. Oxford: Oxford University Press.
- GALLEGO, Ángel/Juan URIAGEREKA (2007) “Conditions on sub-extraction.” In: L. Eguren/O. Fernández-Soriano (eds). *Coreference, modality, and focus*. Amsterdam: John Benjamins, 45–70.
- GRIMSHAW, Jane (1990) *Argument structure*. Cambridge: MIT Press.
- GROHMANN, Kleantes (2003) *Prolific domains*. Amsterdam: John Benjamins.
- HOLES, Clive (2004) *Modern Arabic structures, functions and varieties*. Georgetown University Press: Washington DC.
- HORROCKS, Geoffrey/Melita STAVROU (1987) “Bounding theory and Greek syntax.” *Journal of Linguistics* 23, 79–108.
- ISHIHARA, Shinichiro (2007) “Major phrase, focus intonation, multiple spell-out (MaP, FI, MSO).” *The Linguistic Review* 24, 137–167.
- JOHNSON, Kyle (1991) “Object positions.” *Natural Language and Linguistic Theory* 9, 577–636.
- JOHNSON, Kyle (1992) “Scope and binding theory.” *Syntax and Semantics* 26, 259–275.
- KAHNEMUYIPOUR, Arsalan (2003) “Syntactic categories and Persian stress.” *Natural Language and Linguistic Theory* 12, 333–379.
- KAYNE, Richard (1994) *The antisymmetry of syntax*. Cambridge: MIT Press.
- KHALAILY, Samir (1995) “On the relevance of the split complementizer hypothesis.” In: R. Eckhardt/V. van Geenhoven (eds), *Proceedings of ConSole II*. The Hague: Holland Academic Graphics, 95–111.
- KRATZER, Angelika/Elisabeth SELKIRK (2007) “Phase theory and prosodic spell-out: The case of verbs.” *The Linguistic Review* 24, 93–135.
- LAMBOVA, Mariana (2002) “On A’-movements in Bulgarian and their interaction.” *The Linguistic Review* 18, 327–374.
- LAHNE, Antje (2008) *Where there is fire there is smoke: Local modelling of successive-cyclic movement*. Doctoral dissertation. University of Leipzig.
- LANDAU, Idan (2003) *Modular recoverability: Chain resolution in Hebrew V(P)-fronting*. Ms., Beer-Sheva, Israel.
- LANDAU, Idan (2010) *The locative syntax of experiencers*. Cambridge: MIT Press.
- LASNIK, Howard (1999) *Minimalist analysis*. Oxford: Blackwell.
- LASNIK, Howard (2001) “Subjects, objects, and the EPP.” In: W. D. Davies/S. Dubinsky (eds), *Objects and other subjects: Grammatical functions, functional categories, and configurationality*. Dordrecht: Kluwer, 103–121.
- LEGATE, Julie Anne (2003) “Some interface properties of the phase.” *Linguistic Inquiry* 34, 506–516.
- MCCLOSKEY, James (2000) “Quantifier float and wh-movement in an Irish English.” *Linguistic Inquiry* 31, 57–84.
- MIYAGAWA, Shigeru (2003) “A-movement scrambling and options without optionality.” In: S. Karimi (ed.), *Word order and scrambling*. Oxford: Blackwell, 177–200.



- NESPOR, Marina/Irene VOGEL (1986) *Prosodic phonology*. Dordrecht: Foris.
- NOONAN, Máire (1999) “Successive cyclicity phenomena in wh-movement constructions.” *Proceedings of the West Coast Conference on Formal Linguistics* 18. Somerville: Cascadilla Press, 406–419.
- NUNES, Jairo (2009) “Dummy prepositions and the licensing of null subjects in Brazilian Portuguese.” In: E. O. Aboh et al (eds), *Romance languages and linguistic theory*. Amsterdam: Benjamins, 243–265.
- PESETSKY, David (1995) *Zero syntax*. Cambridge: MIT Press.
- PESETSKY, David (1998) “Some optimality principles of sentence pronunciation.” In: P. Barbosa et al (eds), *Is the Best Good Enough*. Cambridge: MIT Press and MIT WPL, 337–383.
- PESETSKY, David/Esther TORREGO (2001) “T-to-C movement: Causes and consequences.” In: M. Kenstowicz (ed.), *Ken Hale: A life in language*. Cambridge: MIT Press, 355–426.
- PLESSIS, Hans du (1977) “Wh-movement in Afrikaans.” *Linguistic Inquiry* 8, 723–726.
- RACKOWSKI, Andrea/Norvin RICHARDS (2005) “Phase edge and extraction.” *Linguistic Inquiry* 36, 565–599.
- REGLERO, Lara (2007) “Wh-in-situ interrogatives in Spanish.” *Probus* 19, 267–297.
- ROSS, John Robert (1967) *Constraints on variables in syntax*. Doctoral dissertation, MIT, Cambridge.
- SAFIR, Kenneth (1985) *Syntactic chains*. Cambridge: Cambridge University Press.
- SAITO, Mamoru (2009) “Optional A-scrambling.” *Japanese/Korean Linguistics* 16, 44–63.
- SAITO, Mamoru/Keiko MURASUGI (1999) “Subject predication within IP and DP.” In: K. Johnson/I. Roberts (eds), *Beyond principles and parameters: Essays in memory of Osvaldo Jaeggli*. Dordrecht: Kluwer, 167–188.
- SALIH, Mahmud Husein (1985) *Aspects of clause structure in Standard Arabic: A study in relational grammar*. Doctoral dissertation. State University of New York at Buffalo.
- SAWADA, Tsuyoshi (2015) *Pleonastic merger*. Doctoral dissertation. University of Connecticut.
- SCHNEIDER-ZIOGA, Patricia (2009) “Wh-agreement and bounded and unbounded movement.” In: J. Bruccart/A. Gavorro/J. Sola (eds), *Merging features*. Oxford: Oxford University Press, 46–59.
- SELKIRK, Elisabeth (1986) *Phonology and syntax: The relation between sound and structure*. Cambridge: MIT Press.
- SHLONSKY, Ur (1994) “Semitic clitics.” In: M. Starke (ed.), *Geneva Generative Papers* 2. Geneva: University of Geneva, 1–11.
- SHLONSKY, Ur (1997) *Clause structure and word order in Hebrew and Arabic: An essay in comparative Semitic syntax*. Oxford: Oxford University Press.
- SHLONSKY, Ur (2000) “Remarks on the complementizer layer of Standard Arabic.” In: J. Lecarme/J. Lowenstamm/U. Shlonsky (eds), *Research on Afroasiatic Grammar*. Amsterdam: Benjamins, 325–343.

- SIMPSON, Andrew/Zoe WU (2002) “IP-raising, tone sandhi and the creation of S-final particles: Evidence for cyclic Spell-Out.” *Journal of East Asian Linguistics* 11, 67–99.
- STEPANOV, Arthur (2007) “The End of CED?” *Syntax* 10, 80–126.
- STJEPANOVIĆ, Sandra (2003) “A word-order paradox resolved by copy deletion at PF.” *Linguistic Variation Yearbook* 3, 139–177.
- STJEPANOVIĆ, Sandra/Shoichi TAKAHASHI (2001) *Eliminating the Phase Impenetrability Condition*. Presented at the workshop Motivating movement, University of Ulster, January 2001.
- TALIĆ, Aida (2013) “Extraordinary complement extraction.” *Studies in Polish Linguistics* 8, 127–150.
- URIAGEREKA, Juan (1988) *On government*. Doctoral dissertation. University of Connecticut.
- URIAGEREKA, Juan (1999) “Multiple Spell-Out.” In: S. Epstein/N. Hornstein (eds), *Working minimalism*. Cambridge: MIT Press, 251–282.
- VERGNAUD, Jean-Roger (1974) *French relative clauses*. Doctoral dissertation. MIT, Cambridge, Mass.
- WURMBRAND, Susanne (2014) “Tense and aspect in English infinitives.” *Linguistic Inquiry* 45, 403–447.
- ZUBIZARRETA, Maria Luisa (1998) *Prosody, focus, and word order*. Cambridge: MIT Press.
- ZWART, C. Jan-Wouter (1993) *Dutch syntax: A minimalist approach*. Doctoral dissertation. University of Groningen.

### Summary

#### WHAT IS SENT TO SPELL-OUT IS PHASES, NOT PHASAL COMPLEMENTS

An appealing property of the phase theory is that it is relevant to many phenomena, i.e. many domain-based mechanisms are stated in terms of phases. However, although phasal complements have no theoretical status in the phase theory (only phases do), they are taken to define spell-out units. This paper argues for an approach where phases define spell-out domains, which means that what is sent to spell-out is the phase itself. Several arguments to this effect are presented regarding syntax-phonology interaction (in particular, encliticization in Bulgarian and Arabic, stress assignment in German and English, raddoppiamento fonosintattico in Abruzzese, and tone sandhi in Taiwanese), as well as more theoretical issues such as labeling. The assumption, however, has significant consequences for successive-cyclic movement. If phases are sent to spell-out and what is sent to spell-out is inaccessible to the syntax, successive-cyclic movement cannot target phases. Under the account argued for here, successive-cyclic movement therefore does not proceed via phases (i.e. phasal edges). As a result, the account also eliminates the PIC.

**Keywords:** locality of movement, phases, the Phase-Impenetrability Condition, stress assignment, syntax-phonology interface

Povzetek  
NA TOČKI LOČI SE LOČIJO FAZE, NE FAZNA DOPOLNILA

Privlačnost teorije faz leži v dejstvu, da je relevantna pri razlagi več pojavov, pri katerih sodelujejo na domenah osnovani mehanizmi. Čeprav fazna dopolnila nimajo posebnega formalnega statusa v teoriji faz (tega imajo le faze), se uporabljajo kot definicija enote, ki sodeluje pri točki Loči (Spell-Out). V tem članku zagovarjamo pristop, pri katerem so faze tiste, ki definirajo domeno v točki Loči, kar pomeni, da se v točki Loči ločijo faze same in ne njihova dopolnila. V članku je prikazanih več podatkov, ki potrjujejo zagovarjano trditev s področja vmesnika med skladnjo in fonologijo, kot naprimer enklitizacija v bolgarščini in arabščini, naglas v nemščini in angleščini, fonosintaktično podvajanje v abruzijskem narečju in tonemski sandhi v tajvanski mandarinščini. Obravnavanih je tudi več sorodnih teoretičnih vprašanj. Predlagana analiza ima pomembne posledice za zaporedno ciklično premikanje. Če se v točki Loči ločijo faze in če je to, kar se loči, nedostopno za nadaljne postopke v izgradnji, potem zaporedni ciklični premiki ne morejo zadevati faz. Zaporedno ciklično premikanje se torej ne odvija po fazah (tj. faznih robovih). Rezultat analize je tudi izločitev načela o nepredirnosti faz.

**Ključne besede:** lokalnost premika, faze, načelo o nepredirnosti faz, naglas, vmesnik med skladnjo in fonologijo



## DISCONTINUOUS FOCI AND UNALTERNATIVE SEMANTICS

### 1. INTRODUCTION: DISCONTINUOUS FOCI

By the standard question-answer test for focus, we diagnose the VP *was arrested by the police* to be focussed in (1a). Semantically, the same property should be focussed in (1b), but there is no syntactic constituent that corresponds to the putative focus, which consists of the subject and the verb part of the VP. This is what I call a DISCONTINUOUS FOCUS.

- (1) What happened to John?
  - a) John was arrested by the poLICE.
  - b) The police aRRESted John.

If focus is represented by F-markers in the syntax, there are two principled ways of representing discontinuous foci. Either both parts are F-marked, or a constituent dominating both is. These options are illustrated in (2). Note that option number two requires some additional marking – here ‘G’ for ‘given’ –, so as to make sure that the focus is not realized by an accent on the object *John*, the default main stress position in a transitive VP.

- (2) a) [the police]<sub>F</sub> [aRRESted]<sub>F</sub> John  
b) [the police aRRESted John<sub>G</sub>]<sub>F</sub>

Neither solution seems fully satisfactory, as I will show below. (2a) seems incompatible with the prosodic realization of discontinuous foci, because the subject in these constructions is not realized as one would expect from an F-marked constituent; (2b), on the other hand, is semantically a case of overfocussing: it leads to propositional alternatives, rather than property alternatives, as one would expect.

In this paper I propose a new way of handling discontinuous foci, using UNALTERNATIVE SEMANTICS (UAS, see Buring 2015). In UAS, focus is not syntactically marked; rather, focus alternatives are calculated directly from independently needed structural aspects of the representation, in the case at hand: metrical weights. On this view, (1b) naturally comes out as expressing a property-type focus, analogous to (1a), without committing to individual parts of the structure being focused, or allowing for propositional focus alternatives.

---

\* daniel.buring@univie.ac.at

## 2. UNALTERNATIVE SEMANTICS

Like alternative semantics (Rooth 1985, 1992, 1996), UAS systematically maps every syntactic node onto a set of alternatives, i.e., meanings (in the same domain as the node itself) which could be FOCAL TARGETS – for example, the denotation of a previous phrase or utterance with which the sentence containing the focusing is contrasted. Unlike alternative semantics, however, UAS does so by accumulating *restrictions* on focus alternatives bottom to top. Possible focal targets for a constituent A are then those meanings of the same semantic category as that of A which are *not* excluded by these restrictions.

### 2.1 Basics

UAS for English utilizes two essential types of restrictions, weak and strong. A WEAK RESTRICTION is imposed on every branching node that shows default metrical relations, which, for English, are characterized in Figure 1. Weak Restriction excludes all alternatives which result from composing an alternative to the weak sister with the literal meaning of the strong sister. Put differently, it excludes all those alternatives that would be allowed had the weak sister been strong instead.

This is defined in (3a), and abbreviated as in (3b).

(3) WEAK RESTRICTION

If the relative stress among a node  $\alpha$ 's daughters  $S_{\text{strong}}$  and  $S_{\text{weak}}$  accords to the default,  $\alpha$  excludes all focal targets in

- a)  $(\text{alt.dom}(S_{\text{weak}}) \setminus \{\llbracket S_{\text{weak}} \rrbracket_O\}) \otimes_{\alpha} \{\llbracket S_{\text{strong}} \rrbracket_O\}$   
 b)  $x^{S_{\text{weak}}} S_{\text{strong}}$

**alt.dom(S)** is the ALTERNATIVE DOMAIN of S, which I take to be the set of all meanings in the same domain as the meaning of S that can be expressed by expressions of the same category as S.  $\otimes_{\alpha}$  stands for the semantic composition rule needed to combine the alternatives of  $\alpha$ 's daughters. If, for example,  $\alpha$  has daughters  $\beta$  and  $\gamma$ , and the ordinary meaning of  $\alpha$ ,  $\llbracket \alpha \rrbracket_O$  is  $\llbracket \beta \rrbracket_O \llbracket \gamma \rrbracket_O$ ; then  $\text{alt.dom}(\beta) \otimes_{\alpha} \text{alt.dom}(\gamma)$  would be  $\{b(g) \mid b \in \text{alt.dom}(\beta) \& g \in \text{alt.dom}(\gamma)\}$ .

Another way of thinking of Weak Restriction is that it says ‘if the weak sister is (part of the) focus, so is the strong sister’, or: ‘the weak sister is not a narrow focus’. The excluded meanings – e.g., those characterized by (3a)/(3b) – are generally also written as  $\llbracket \alpha \rrbracket_U$ .

The second type of restriction, STRONG RESTRICTION, applies where the metrical weights between two sister nodes have been REVERSED. In this case the restriction requires that the metrically PROMOTED sister be part of a focus. This is formalized in (4).

WEAK	STRONG
functional	lexical
head	complement
left projection	right projection

Figure 1: Structural metrical defaults, in descending order of importance

4) STRONG RESTRICTION

If  $a$  has undergone prosodic reversal,  $a$  excludes all focal targets in  $alt.dom(a) \setminus ((alt.dom(S_{strong}) \setminus \llbracket S_{strong} \rrbracket_O) \otimes_a \{ \llbracket S_{weak} \rrbracket_O \})$

Another way of saying this is ‘strong daughter is a focus, weak daughter is not’. It is convenient to characterize the restrictions imposed by Strong Restriction in terms of the alternatives it *allows*, rather than the excluded ones. This is symbolized as in (5) (rather than  $P^{x \setminus S_{strong} S_{weak}}$ ). The reader should bear in mind, though, that technically all restrictions are in the form of sets of unalternatives, as defined in (4).

(5)  $x \setminus S_{strong} S_{weak}$

The workings of Weak and Strong Restriction are illustrated in Figures 2 and 3. In addition to the two restrictions, UAS requires a mechanism of unalternative PROPAGATION. Defined as in (6) – where  $\llbracket S \rrbracket_u$  stands for the UNALTERNATIVES of S, i.e., the meanings excluded as FTs by Weak and Strong Restriction – Propagation makes sure that (weak or strong) restrictions introduced on lower nodes are propagated all the way up to the root node.

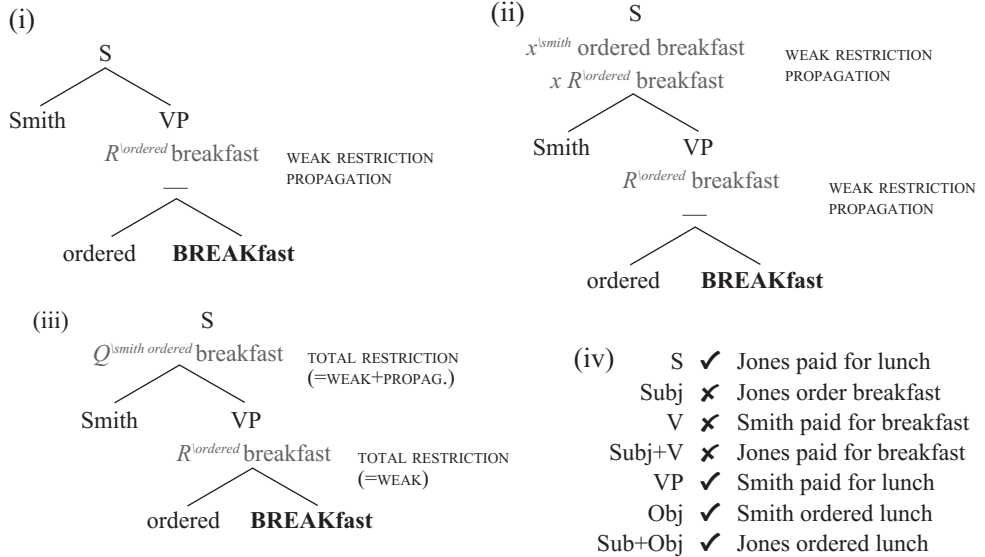


Figure 2: Unalternatives calculated at the VP, (i), and S level, (ii); (iii) shows how the (weak) restriction introduced at the S level, and the weak restriction propagated from the VP level can be combined into a single restriction. The resulting predictions regarding what can and cannot be a focal target according to  $Q^{smith ordered}$  breakfast are shown in the table in (iv). The leftmost column indicates what kind of focus this would correspond to in a system using F-marking.

(6) PROPAGATION

any branching node  $[\alpha S_1 S_2]$  excludes all focal targets in  $(alt.dom(S_1) \otimes_{\alpha} \llbracket S_2 \rrbracket_{\mu})$  as well as those in  $(\llbracket S_1 \rrbracket_{\mu} \otimes_{\alpha} alt.dom(S_2))$

In most cases restrictions imposed by strong and weak restriction and by propagation can be combined into a single restriction, as in Figures 2 and 3.

The overall set of possible Focal Targets for the two structures in Figures 2 and 3 are summarized in the tables in Figure 2(iv) and Figure 4.

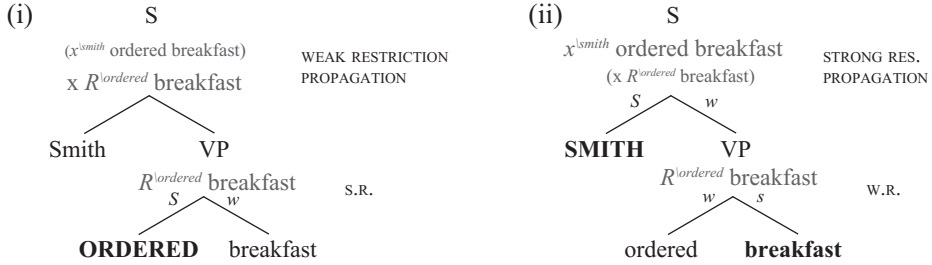


Figure 3: Strong restriction, applied at the VP- and S-levels (3(i)/3(ii), respectively); note that in the latter case, normal metrical weights,  $w$ - $s$ , are assigned within the VP.

<i>Smith ORDERED breakfast</i> $x R^{ordered} \text{ breakfast}$	<i>SMITH ordered breakfast</i> $x^{smith} \text{ ordered breakfast}$
S ✗ Jones paid for lunch	S ✗
Subj ✗ Jones order breakfast	Subj ✓
V ✓ Smith paid for breakfast	V ✗
Subj+V ✓ Jones paid for breakfast	Subj+V ✗
VP ✗ Smith paid for lunch	VP ✗
Obj ✗ Smith ordered lunch	Obj ✗
Sub+Obj ✗ Jones ordered lunch	Sub+Obj ✗

Figure 4: Predictions for prosodic reversal at the VP and S level, respectively.

## 2.2 Discontinuous Foci in Unalternative Semantics

The tables in Figure 2(iv) and in Figure 4 each contain one licit discontinuous focus, Subj+Obj and Subj+V, respectively. The latter of these is the kind we looked at in section 1, transitive subject plus verb. This is illustrated in detail in Figure 5(i). Figure 5(ii) schematizes the configuration in which discontinuous foci can come about in Unalternative Semantics in general: whenever the strong daughter  $\alpha$  of a branching node  $\mu$  with

default stress among its daughters, is itself branching; in that case  $\alpha$ 's strong daughter,  $C$ , and its 'aunt',  $A$  may be interpreted as a discontinuous focus.

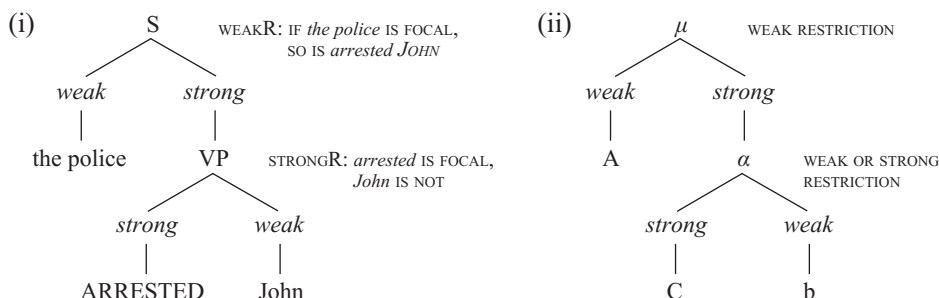


Figure 5: The restrictions imposed on 5(i), the tree corresponding to (1b), are compatible with *the police arrested* (but not *John*) being focal, so that ‘what happened to John’, or the meaning of any answer to it, is a possible Focal Target. Generally, UAS predicts that two constituents  $A+C$  can be interpreted as a discontinuous focus in exactly two configurations, those in 5(ii), with the order of sisters irrelevant. Note that if has default metrical weights (and hence Weak Restriction applies), the entire discontinuous focus structure will be realized by default stress.

### 3. OTHER INSTANCES OF DISCONTINUOUS FOCI

Using the standard question/answer test for pragmatically determining the focus in an answer, we can actually rather frequently diagnosed a discontinuous focus.

For example, (7a) shows a discontinuous focus consisting of a transitive verb and an indirect object to the exclusion of the direct object *the books*. Note that (7b) actually displays the same kind of discontinuous focus, except that in this case no prosodic reversal (i.e., deaccenting) has taken place – so this case is less spectacular on the surface.

- (7) What did you do with the books?  
 a) I sent KIM<sub>F</sub> the books.  
 b) I sent the books to KIM<sub>F</sub>.

As anticipated in the table in Figure (2iv), a discontinuous focus may also consist of the subject and the object of a transitive sentence, as in (8).

- (8) (A lot of people were introduced to the mayor, but) no-one/nun introduced SUE<sub>F</sub>.

(8) is another case in which DF is realized by default prosody. We can, however, change it so as to make sure that indeed only the object, but not the verb, is part of the



focus. The key to that is the first of the prosodic defaults in Figure 1: between a functional and a lexical sister, the lexical one is strong by default, regardless of linear order or other factors. This can be seen in the broad focus examples in (9).

- (9) If Sue continues to work like this, John will  
a) ... HIRE her/ introDUCE himself/ SAY something/ REPORT that.  
b) ... hire HELP/ introduce new RULES/ say the TRUTH/ report the OPposite.

If we now embed a VP of this sort – strong verb, weak pronominal object – in a configuration like (8), we get a non-default stress pattern:

- (10) (A lot of people were introduced to the mayor, but) no-one/nun introduced ME<sub>F</sub>.

The default pattern in this case would accent the verb instead, as in (11).

- (11) (A lot of people were introduced to the mayor, but ...) no-one/nun TOLD me<sub>F</sub>.

So we can indeed be sure that (10) is an instance of discontinuous focus, and, by parity of reasoning, (8) is, too. I will return to the significance of cases like these in section 4.2.

Indeed, such discontinuous foci can even result in both ‘extraordinary accenting’ (of a pronoun) *and* deaccenting at the same time.

- (12) (A lot of people were introduced to the mayor, but) no-one/nun introduced ME<sub>F</sub> to the mayor.

All of these examples also illustrate that prosodic reversal (i.e., non-default stress) is not the same deaccenting; prosodic reversal in (9a) and (10) manifests itself in the *addition* of a pitch accent on the functional object *me*, which would otherwise be unstressed. The same effects can be observed in (13), where an additional stress, which happens to end up being the nuclear stress of the sentence, is put on the verbal particle *back* (the normal intonation here would have the nuclear pitch accent on *books*).

- (13) (What happened to the books?) Sam sent the books BACK<sub>F</sub>.

Accent addition as a result of prosodic reversal is also commonly found in head final verb phrases, where the second default from Table 1 – that a predicate is weak and its argument strong – applies. Dutch and German VPs provide a well-known case in question, as in example (14). Here too, neutral stress would fall on the object *Bücher*, rather than the sentence-final verb. Again prosodic reversal has taken place, resulting in an additional pitch accent on the structurally weak head-sister of the phrase.

- (14) Where are the books?  
 KIM hat die Bücher ABgeholt<sub>F</sub>.  
 K. *has the books up picked*  
 ‘Kim picked the books up.’

I mention these different instances of discontinuous focus because they, collectively, pose severe difficulties for alternative representations, which I will discuss in the next section.

#### 4. A LOOK AT ALTERNATIVE ANALYTICAL OPTIONS

In this section we will take a closer look at alternative analyses for the phenomenon of discontinuous focus. As said at the outset, it would seem that the two major analytic options are to analyze discontinuous focus as an instance of multiple focus, or as an instance of broad focus with focus-internal givenness deaccenting. We will inspect these options in turn.

##### 4.1 Discontinuous Focus vs. Double Focus

(15a) illustrates what the initial example would be analyzed like using F-markers on the individual parts of the discontinuous focus (whereas (15b) indicates the kind of analysis advocated here, though recall that no actual syntactic marking of the focused constituents is part of the theory). Standard alternative semantics would actually assign the same set of focus alternatives to this structure that it would to an ordinary VP focus like the one in (1a).

- (15) claim:  
 a) \*[the police]<sub>F</sub> [arrested]<sub>F</sub> John  
 b) the police arrested<sub>F</sub> John

So while semantically no argument against (15a) would be forthcoming, we may object to the predictions such an analysis would make regarding the prosodic realization of such examples. Notice that according to the F-marking pattern in (15), both the subject *the police* and the transitive verb *arrested* are foci in the sense of for example Selkirk (1995) (there called FOCus). That is to say, both are F-marked constituents which are not dominated by another F-marker.

One prediction of this representation is that both parts of the discontinuous focus should be realized as foci; in particular neither of them should be able to undergo phonological reduction. The examples in (16), however, clearly show that this prediction is not born out. The subjects in (16a) and (16b) may be reduced to the unstressed forms *smon* and *ya*. In (16c) it is even possible to drop the subject pronoun altogether.

- (16) Where’s Kim’s homework?/What happened to Kim’s homework?  
 a) Someone/Smone STOLE it.

- b) You're/Ya SITTING on it.
- c) I/∅ forgot it on the BUS.

Needless to say, none of this is expected if the subjects were themselves foci.<sup>1</sup> This expectation is confirmed in the examples in (17) which are indeed double foci. In these cases reduction of the subject pronouns – either segmentally or by deletion – is completely impossible.

- (17) a) (First, JO STOLE my homework and now) YOU are/\*ya SITTING on it.
- b) (First my brother misplaced my homework and now what happened?!) \*(I forgot it on the BUS.

A second argument in the same vein involves subject+transitive-verb foci which are realized by nuclear stress on the subject, as in (18a). However, this argument is currently more suggestive than decisive, since I do not fully understand the factors that lead to this configuration in the first place.

- (18) John looks devastated. What happened to him?
- a) Maybe his WIFE left him.
- b) Maybe his WIFE died.
- c) Maybe his DOG ran away.

If *left* in (18a) were itself a focus (not just part of one), it should not be able to remain unaccented. Standard cases in which focal predicates remain unaccented always involve them forming *one* larger focus with an argument, as would be the case if *his wife left* were a discontinuous focus in (18a). In that sense, (18a) is a problem for any analysis that assigns it the focus structure [*his wife*]<sub>F</sub>[[*left*]<sub>F</sub> *him*].

In most cases discussed in the literature, this kind of INTEGRATION – realizing a broad V+argument focus with just one accent on the argument – involves internal arguments, including possibly unaccusative subjects as in (18b). Unergative intransitive subjects as in (18c) do partake in this pattern, too, as observed for example in Krifka (1984). Realizing a broad focus on a transitive subject, as in (18a), however, is generally thought to be impossible, and consequently excluded by standard algorithms for realizing focus, including the one assumed here. Evidently, this is in need of refinement: at least in some cases, including (18a), even transitive subjects may realize a broad focus, provided the internal argument can remain accent-less for independent

---

1 An anonymous reviewer suggests that, alternatively, ‘weak pronouns, weak indefinites and the likes’ should never be in the focus (so that (16) would be narrow V focus), or that, yet alternatively, the non-accenting of the subjects in (16) ‘can be handled by the PF, through some kind of algorithm which reduces the prosodic shape in a particular set of cases’ (hence suspending the assumption that a focus must contain an accent). Either move would require significant modifications to focus pragmatics and focus realization, respectively. Since I am not aware of any proposals with such features, I cannot pursue these lines of analysis further here.

reasons, like *him* in (18a).<sup>2</sup> Whatever the details, the accent pattern in (18a) is hard to reconcile with a representation that has both V and the subject be foci.

## 4.2 Broad Focus with Anaphoric Deaccenting?

Let us then turn to the second analytical possibility: could these examples be actually instances of broad, e.g., sentential, foci, within which anaphoric deaccenting has taken place?

The only standard test for teasing apart broad foci with deaccenting within them and narrower foci that I am aware of is ellipsis. Thus in (19), it is impossible to elide the verb phrase, even though *will resign* can be deaccented as given, and even though we know independently that bare subject answers are possible in English, if the subject is a narrow focus.

- (E) (What will happen if Sam resigns?)
  - a) (Then) KIM will resign.
  - b) #KIM.
  
- (E') (Who will oversee the project if Sam resigns?)
  - a) (Then) KIM will oversee the project.
  - b) KIM.

Unfortunately the ellipsis test is, as far as I can see, not applicable for the cases of discontinuous focus we are interested in here. This is probably due to the fact that the remnant of ellipsis itself has to be a constituent, which would mean that omission of the background in a discontinuous focus structure is by definition impossible. What we need, then, is a new diagnostic for the difference between deaccenting within a focus and discontinuous focus. In what follows I will tentatively investigate two avenues towards probing this difference in sections 4.2.1 and 4.2.3.

### 4.2.1 Pragmatic Contrast

My first attempt at mounting argument against the broad-focus-plus-deaccenting analysis of discontinuous foci involves the size of the focus, as probed by pragmatic contrast,

---

2 As with intransitive subject integration, such cases characteristically involve Allerton and Crutenden's (1979) verbs of appearance and disappearance, and 'verbs expressing misfortune' (ostensibly what we have in (18)). An anonymous reviewer suggests that the verbs in (18) are 'already implied by the ... subject' and hence not in need of focussing. However, as the reviewer themselves points out, the fact that different predicates would have to be implied by the same subject in the same context in (a) vs. (b) casts doubt on such a story. It may also be of interest to note that a parallel context implicating fortunate events, rather than misfortune, does not seem to allow for integration in the same way:

- (i) John looks happy. What happened?
  - a) Maybe his PARTNER wants to marry him.
  - b) #Maybe his DAUGHTER graduated.

as well as the focus sensitive particle *also*. What is crucial to the argument is the fact that deaccenting within a focus does not affect the focus alternatives generated by standard alternative semantics. For example, if the second clause in (19) were to be represented as in (20a), it should permit verb phrase focus alternatives, that is: properties.

(19) What did the guy who delivered the fridge do wrong? He scratched the EDges of the fridge.

(20) claim:

- a) \*he [scratched the EDges of [the fridge]<sub>G</sub>]<sub>F</sub>.
- b) he scratched the EDges of<sub>F</sub> the fridge.

According to the analysis advocated in the present paper, on the other hand, the structure is more akin to that in (20b), in which the noun phrase *the fridge* is not part of the focus and therefore not able to license focus alternatives. According to that analysis, only focal targets that describe relations between the delivery person and the fridge are permitted by (19).

This I attempt to put to the test in (21). We observe that (21a) sounds odd in the context given. This is expected if – as assumed here – the focus excludes (*of*) *the fridge* and thus the alternatives available to *also* are of the form ‘they did x to the fridge’, since the context specifies no other damage done to the fridge. In (21b), on the other hand, we clearly have VP focus, permitting alternatives ‘they Q’, so that ‘they broke the glass door’ can satisfy the presuppositions of *also*.

(21) So ok, the guys who delivered the fridge broke the French doors in the living room. What other damage did they do? ...

- a) #They also scratched the EDges of the fridge.
- b) They also scratched the edges of the FRIDGE.

According to the ‘broad-focus-plus-deaccenting’ view, (21a) and (21b) have the exact same F-pattern (the sole difference being G-marking on the fridge) and focus alternatives, (22).

- (22) a) they also [scratched the edges of [the fridge]<sub>(G)</sub>]<sub>F</sub>
- b) ‘they also Q’

So the felicity of *also* should be unaffected by whether or not *fridge* is deaccented. In fact, if anything, (21a) should be the preferred realization, given the generally held view that givenness marking should be maximized.<sup>3</sup> This strongly suggests that *the*

3 E.g., ‘Do Not Overlook Anaphoric Possibilities’ in Williams (1997), or analogous principles that minimize focus marking, e.g., ‘AvoidF’ in Schwarzschild (1999), the requirement that given elements must at most be FOCUS marked in Selkirk (1995), ‘Maximize Presupposition’ in Sauerland (2005), or ‘Maximize Background’ in Büring (2012).

*fridge* is not part of the focus in (21a), i.e., that we are dealing with a discontinuous focus, *scratched the edges of*.

I used *also* in (21) to check the maximum size of a focus: the more focus alternatives are available, the more permissive (regarding possible antecedents) *also* should become. A parallel argument can be mounted using free focus. The contrast in (23) exactly parallels that in (21).

- (23) So ok, the guys who delivered the fridge broke the French doors in the living room. What other damage did they do? ...
- a) #They scratched the EDges of the fridge.
  - b) They scratched the edges of the FRIDGE.

As before, if ‘broke the French doors in the living room’ were among the alternatives to *scratched the EDGES of the fridge* (in this context, that is, where *fridge* is given) (23a) should be as good as, if not better than, (23b). In contradistinction, the discontinuous focus analysis favored here again correctly predicts that only (23b) will be felicitous here, if ‘broke the glass door’ is to be the focal target.

These examples have implication beyond determining the size of the focus, which I will only touch upon here. They suggest that, in general, there is no such thing as pure givenness deaccenting within a broad focus; this implies in turn that, for example, the focal targets in (19) must be (other) ‘things he did to the fridge’, not ‘things he did wrong’ in general, even though that is not, strictly speaking, salient in the context. (21) and (23) furthermore indicate that ‘deliver the fridge’ is not an available focal target for *scratch the EDGES of the fridge* in these examples (otherwise (21a) and (23a) should be fine, as ‘things they do with/to the fridge’ are permitted alternatives, even under the discontinuous focus analysis); ostensibly this is because there is no pragmatic contrast between delivering the fridge and scratching its edges; a full account of focussing needs to model such a contrast requirement, which e.g., the proposals in Rooth (1992) and Schwarzschild (1999) do not (see Wagner 2006b, 2012 and Katzir 2013 for discussion, Büring (forthcoming) for a recent proposal utilizing Unalternative Semantics).

The basic contrast used to argue against the ‘broad-focus-plus-deaccenting’ analysis of discontinuous foci, however, is, I think, not affected by these considerations; the crucial fact is simply that these cases do not behave like VP- or sentential foci.

#### 4.2.2 Excursus: Broad Focus with Givenness Movement

An anonymous reviewer drew my attention to an alternative version of the broad-focus-plus-deaccenting analysis, on which the given/deaccented elements within the focus are moved out of the focus, either string vacuously, where applicable, or at ‘Logical Form’, as in (24) (thanks to the reviewer, who also suggested the structures on which those in (24) are based).

- (24) a) John<sub>1</sub> [the police arrested t<sub>1</sub>]<sub>F</sub>  
 b) I<sub>2</sub> [the books<sub>1</sub> [t<sub>2</sub> sent Kim t<sub>1</sub>]<sub>F</sub>]



- |   |     |
|---|-----|
| a) Marie přiměla k odchodu <u>Jiřího</u> <sub>F</sub> . | 7.9 |
| <i>M.NOM convinced to leaving J.ACC</i>                 |     |
| ‘Marie convinced Jiri to leave.’                        |     |
| b) Marie přiměla <b>Jiřího</b> <sub>F</sub> k odchodu.  | 7.3 |

The mean ratings of 7.9 and 7.3 do not, according to Groeben et al. (to appear) show a significant preference for either of these options. Stress-shift *within* a focus, on the other hand, is dispreferred; reordering is preferred instead.

- |   |             |
|---|-------------|
| (27) Q: Nevíš, jestli už všichni odešli?                  | mean rating |
| ‘Do you have an idea if everyone left yet?’               |             |
| a) <u>Marie přiměla k odchodu Jiřího</u> <sub>F</sub> .   | 5.9         |
| <i>M.NOM convinced to leaving J.ACC</i>                   |             |
| ‘Marie convinced Jiri to leave.’                          |             |
| b) # <u>Marie přiměla Jiřího</u> k odchodu <sub>F</sub> . | 4.9         |

Here the difference in mean acceptability ratings is significant: reordering is the preferred realization. We should now be able to use the different preference patterns to probe whether a given construction shows narrow, possibly discontinuous, focus, or broad focus with internal deaccenting. Groeben et al.’s (to appear) own data may actually provide a case in question: in certain conditions, the expected preference for reordering was not found, see (28).

- |   |             |
|---|-------------|
| (28) Q: Vyzvala Marie Jiřího k odchodu?                 | mean rating |
| ‘Did Marie ask Jiri to leave?’                          |             |
| a) Marie Jiřího k odchodu <b>přiměla</b> <sub>F</sub> . | 7.1         |
| <i>M.NOM J.ACC to leaving convinced</i>                 |             |
| ‘Marie convinced Jiri to leave.’                        |             |
| b) Marie <b>přiměla</b> <sub>F</sub> Jiřího k odchodu   | 7.2         |
| (29) Q: Nevíš, proč Jiří odešel?                        | mean rating |
| ‘Do you have an idea why Jiri left?’                    |             |
| a) <u>Marie Jiřího k odchodu přiměla</u> <sub>F</sub> . | 6.7         |
| <i>M.NOM J.ACC to leaving convinced</i>                 |             |
| ‘Marie convinced Jiri to leave.’                        |             |
| b) <u>Marie přiměla Jiřího k odchodu</u> <sub>F</sub> . | 6.3         |

Groeben et al. (to appear) speculate that (28b) and (29b) may involve ‘focus accommodation’ (to a narrow V focus). But this may in fact be a case of discontinuous focus, which would explain the lack of an effect: (29) *is* an instance of ‘moving background to the left of focus’, as indicated in (30) (as before, keep in mind that the F-marking is for perspicuity only).



- (30) Q: Neviš, proč Jiří odešel?  
 ‘Do you have an idea why Jir left?’
- a) Marie Jiřího k odchodu **přiměla**<sub>F</sub>  
*M.NOM J.ACC to leaving convinced*  
 ‘Marie convinced Jiří to leave.’
- b) Marie **přiměla**<sub>F</sub> Jiřího k odchodu.

Note that this argument hinges on the premise that in Czech, *both* discontinuous focus and focus-internal deaccenting, are possible – possibly unlike in English, if our remarks in the previous subsection were on the right track. I will leave it for future work to explore these differences in more depth.

## 5. SUMMARY

In this paper I have demonstrated how Unalternative Semantics provides a natural way to model discontinuous foci. Several cases from English were discussed which should plausibly be analyzed as such. Crucially, since Unalternative Semantics does not use syntactic F-markers to represent focus, the questions whether ‘the’ focus is a constituent, as well as whether a discontinuous focus is the same as multiple foci, cannot even arise.<sup>5</sup> As discussed above, either way of answering it in a framework that uses F-marking leads to problematic consequences for either focus realization or interpretation, indirectly arguing the case for the type of treatment advocated here.

## References

- ALLERTON, D.J./A. CRUTTENDEN (1979) “Three Reasons for Accenting a Definite Subject.” *Journal of Linguistics* 15/1, 49–53.
- BÜRING, Daniel (2012) “What’s New (and What’s Given) in the Theory of Focus?” In: Sarah Berson/Alex Bratkievich/Daniel Bruhn/Ramon Escamilla/Amy Campbell/Allegra Giovine/Lindsey Newbold/ Marta Piqueras-Brunet/Marilola Perez/ Russell Rhomieux (eds), *Proceedings of the Thirty-Fourth Annual Meeting of the Berkeley Linguistics Society*, February 8-10 2008. Berkeley, CA: Berkeley Linguistics Society, 403–424.
- BÜRING, Daniel (2015) “Unalternative Semantics.” In: Sarah D’Antonio/Mary Moroney/Carol Rose Little (eds), *Proceedings of SALT 25*. Washington: Linguistic Society of America, 550–575.
- BÜRING, Daniel (2016) *Intonation and Meaning. Surveys in Semantics and Pragmatics*. Oxford: Oxford University Press.

5 I do think that, alongside the discontinuous focus pattern discussed here, a true double focus pattern exists, which, among other things, realizes both foci as intermediate phrase heads, and thus prohibits prosodic reduction of both focused elements. A tentative analysis for these cases is offered in Büring (forthcoming).

- BÜRING, Daniel (forthcoming) “Focus, Questions and Givenness.” In: Klaus von Heusinger/Edgar Onea/Malte Zimmermann (eds), *Questions in Discourse*. Holland: Brill.
- GROEBEN, Lena/Radek ŠÍMÍK/Frank KÜGLER (to appear) “Stress Shift, Focus, and Givenness in Czech.” In: *Proceedings of Formal Approaches to Slavic Linguistics (FASL) 24*. Michigan Slavic Publications.
- KATZIR, Roni (2013) “A Note on Contrast.” *Natural Language Semantics* 21/4, 333–343.
- KRIFKA, Manfred (1984) *Fokus, Topik, syntaktische Struktur und semantische Interpretation*. Available: <http://amor.cms.hu-berlin.de/~h2816i3x/Publications/Krifka%201984%20Fokus.PDF>.
- KRIFKA, Manfred (1993) “Focus and Presupposition in Dynamic Interpretation.” *Journal of Semantics* 10, 269–300.
- KUČEROVÁ, Ivona (2007) *The Syntax of Givenness*. PhD dissertation. MIT.
- ROOTH, Mats (1985) *Association with Focus*. PhD dissertation. UMass Amherst.
- ROOTH, Mats (1992) “A Theory of Focus Interpretation.” *Natural Language Semantics* 1, 75–116.
- ROOTH, Mats (1996) “Focus.” In: Shalom Lappin (ed.), *The Handbook of Contemporary Semantic Theory*. London: Blackwell, 271–297.
- SAUERLAND, Uli (2005) “Don’t interpret focus: Why a presuppositional account of focus fails, and how a presuppositional account of givenness works.” In: *Proceedings of Sinn und Bedeutung* 9. Nijmegen, 370–384.
- SCHWARZSCHILD, Roger (1999) “GIVENness, AvoidF and Other Constraints on the Placement of Accent.” *Natural Language Semantics* 7/2, 141–177.
- SELKIRK, Elisabeth (1995) “Sentence Prosody: Intonation, Stress, and Phrasing.” In: John A. Goldsmith (ed.), *The Handbook of Phonological Theory*. London: Blackwell, 550–569.
- ŠÍMÍK, Radek/Marta WIERZBA (2015) “The Role of Givenness, Presupposition, and Prosody in Czech Word Order: An Experimental Study.” *Semantics and Pragmatics* 8/3, 1–103.
- WAGNER, Michael (2005) *Prosody and Recursion*. PhD dissertation, MIT.
- WAGNER, Michael (2006a) “Association by movement: evidence from NPI-licensing.” *Natural Language Semantics* 14/4, 297–324.
- WAGNER, Michael (2006b) “Givenness and Locality.” *Semantics and Linguistic Theory (SALT)* 16, 295–312.
- WAGNER, Michael (2010) “Prosody and Recursion in Coordinate Structures and Beyond.” *Natural Language and Linguistic Theory* 28, 183–237.
- WAGNER, Michael (2012) “Focus and Givenness: A Unified Approach.” In: Ivona Kučerová/Ad Neeleman (eds), *Contrasts and Positions in Information Structure*. Cambridge: Cambridge University Press, 102–147.
- WILLIAMS, Edwin (1997) “Blocking and Anaphora.” *Linguistics Inquiry* 28, 577–628.

Summary  
DISCONTINUOUS FOCI AND UNALTERNATIVE SEMANTICS

Discontinuous foci – cases in which the focus as expected by semantic or pragmatic considerations is not a single constituent within the phrase marker – are not commonly discussed in the formal literature on focussing. This paper proposes to use Unalternative Semantics to analyze such foci. Unalternative Semantics is a novel framework for calculating focus alternatives from metrically annotated trees (instead of trees with F-makers); this format naturally lends itself to the modelling of discontinuous foci. The paper compares this approach to other, alternative options involving F-markers and argues in favor of the F-less treatment.

**Keywords:** focus, discontinuous focus, alternative semantics, deaccenting, givenness

Povzetek  
PREKINJENI FOKUSI IN NEALTERNATIVNA SEMANTIKA

Prekinjeni fokusi – primeri, v katerih fokus, določen glede na semantične in pragmatične značilnosti, ni enoten sestavnik – so v literaturi o fokusu redkeje obravnavani. Pričujoči članek pokaže razčlenbo takšnih zgradb s pomočjo nealternativne semantike. Nealternativna semantika je novejši okvir, s katerim preračunavamo fokusne alternative iz metrično označenih dreves (namesto iz dreves označenih s F-oznaki); takšen pristop je posebej primeren za modeliranje prekinjenih fokusov. Članek primerja ta pristop z drugimi, alternativnimi možnostmi, ki vključujejo F-oznake, in zagovarja razčlenbo brez takšnih oznak.

**Keywords:** fokus, prekinjeni fokus, alternativna semantika, nenaglašenost, danost



## A SOURCE OF PARAMETRIC VARIATION IN THE LEXICON

### 1. INTRODUCTION

An influential conjecture concerning parameters is that they can possibly be “restricted to formal features [of the lexicon]” (Chomsky 1995: 6; cf. Borer 1984: 2f).

In Rizzi (2009, 2011) such features are understood as instructions triggering one of the following syntactic actions:

- (1) a) External Merge
- b) Internal Merge (Move)
- c) Pronunciation/Non pronunciation (the latter arguably dependent on Internal Merge, Kayne 2005a)

Here I discuss what appears to be a particularly pervasive source of variation among languages in the domain of the lexicon (both functional and substantive) and consider whether and how it can be reduced to one of the above actions.

The variation can be preliminarily characterized as follows: **language A has two (or more) lexical items which correspond to just one lexical item in language B.**

### 2. FUNCTIONAL LEXICON

#### Example 1 (Zanuttini 1997: §3.3.1 and §3.3.2)

The Piedmontese northern Italian dialect of Turin has two sentential negative markers: *nen*, which is a neutral negative marker (it simply negates a certain proposition P), and *pa*, corresponding to standard Italian *mica*, which is a presuppositional negative marker (it negates a certain proposition P that the speaker believes to be presupposed in the context of utterance, with the effect of denying the correctness of such a presupposition).<sup>1</sup>

The northern Italian Valdotain dialect of Cogne, on the other hand, has only one sentential negative marker, *pa*, which covers both functions (it can be used either as a neutral or as a presuppositional negative marker).

\* cinque@unive.it

\*\* I wish to thank for their comments to an oral presentation of this work Richard Kayne, Iliyana Krapova, Rita Manzini, Henk van Riemsdijk, Luigi Rizzi, and Peter Svenonius, as well as two anonymous reviewers of *Linguistica* for their comments.

1 For a discussion of the pragmatic conditions required for such presuppositional negation to be felicitous see Cinque (1976) and Zanuttini (1997, Chapter 3).

### Example 2 (Cinque 1999: §4.19, §4.25 and 208 n57)

The two English adverbs *soon* and *early* seen in (2a) and (2b) are rendered in Italian by the single adverb *presto*, as seen in (3a-b):<sup>2</sup>

- (2) a) *He will soon have to get up.*  
b) *He will have to get up early.*
- (3) a) *Presto si dovrà alzare.*  
'He will soon have to get up.'  
b) *Si dovrà alzare presto.*  
'He will have to get up early.'

## 3. SUBSTANTIVE LEXICON

### Example 1:

While Italian has separate lexical items to refer to 'arm' and 'hand', *braccio* and *mano* respectively, 'leg' and 'foot', *gamba* and *piede* respectively, Bulgarian uses one lexical item for both 'arm' and 'hand', *raka*,<sup>3</sup> and one lexical item for both 'leg' and 'foot', *krak*.

### Example 2

To the distinct English lexical items *grandson/granddaughter* (i.e. male/female grandchild) and *nephew/niece* (i.e. male/female child of sibling), only one lexical item corresponds in Italian: *nipote*, for 'grandson'/'granddaughter'/'nephew'/'niece'.

The examples could easily be multiplied.

## 4. THE LOGIC UNDERLYING THIS PATTERN OF VARIATION

I take this pattern of variation not to be accidental, and to arise from the fact that the functional or substantive denotata of the two (or more) lexical items of language A which correspond to the unique lexical item of language B **share one component/feature** (while differing with respect to other components/features). Language B capitalizes on this shared component/feature. That is, the single lexical item of language B is uniquely specified for the common component/feature and left unspecified (in ways to which I return) for the differentiating components/features. Language A, on the other

---

2 I take these adverbs (more exactly, adverbial phrases) to belong to the functional lexicon as they appear to correspond in terms of position and interpretation to two independent aspectual projections (cf. Cinque 1999, Chapter 3).

3 For the lexicalization of 'hand' and 'arm' cross-linguistically the World Atlas of Language Structures (<http://wals.info/chapter/129>) reports that 228 languages have an identical word and 389 languages have two different words.

hand, capitalizes on the differentiating components/features. That is to say, its two (or more) lexical items corresponding to the single lexical item of language B are specified both for the shared component/feature and for the differentiating component(s)/feature(s).

Consider in this light the examples of the functional lexicon given in section 2 (§4.1) and those of the substantive lexicon given in section 3 (§4.2).

#### 4.1 The Functional Lexicon

Concerning the first example of section 2, we noted there, after Zanuttini (1997, §3.3.1 and §3.3.2), that the Piedmontese of Turin has two sentential negative markers, the presuppositional *pa* and the neutral *nen*. The two, in addition to the different pragmatic conditions that govern them, also differ in the position they occupy within the clause. The presuppositional negative marker *pa* necessarily precedes an adverb like *gia* ‘already’, while the neutral negative marker *nen* necessarily follows it:

- (4) a) *A l'è pa/\*nen gia parti.*  
 Cl Cl is neg already left  
 ‘He hasn’t already left.’
- b) *A l'avia gia nen/\*pa salutami cul di la.*  
 Cl Cl had already neg greeted-me that day there  
 ‘Already on that day he had not greeted me’

The overall order given in Zanuttini (1997,72) is thus **pa** > **gia** > **nen** (> *sempre*). In fact, the two can co-occur, in the expected order (*pa nen*), as shown in (5):

- (5) *Fa pa nen (\*nen pa) suli.*  
 do neg neg that  
 ‘Don’t do that!’ (the assumption is that the addressee is about to do it)

The Valdotain dialect of Cogne, on the other hand, has only one negative marker, which can be used either as a presuppositional or as a neutral negation: *pa*. However, this is not merely a lexical quirk. When it is presuppositional *pa* precedes *dza* ‘already’; when it is neutral it follows *dza*:

- (6) a) *L'è pa dza parti?*  
 Cl is neg already left  
 ‘He hasn’t already left, has he?’
- b) *I m'a dza pa saluià ce dzor lai.*  
 Cl me has already neg greeted that day there  
 ‘Already that day he didn’t greet me.’

The overall order given in Zanuttini (1997,82) is thus **pa** > **dza** > **pa** (> *toujou*).<sup>4</sup>

All of this seems to me to point to the presence of two specialized negative projections, which share a common core (**Negation of P**); one below the projection occupied by the adverb ‘already’, expressing simple Negation of P, and one above it, expressing Negation of P, where P is presupposed (denial of P). We can assume that these functional projections (like any other such projection) are built by selecting the relevant feature/head from the functional lexicon according to the selection properties of that feature.

If the lexical specifications of Piedmontese *nen* and *pa* are [Negation of P] and [Negation of P, P presupposed], respectively, each will be uniquely matched with the corresponding projection. If on the other hand the lexical specification of Valdotain *pa* is [Negation of P], with unspecified [P presupposed], then it will be able to match either projection.<sup>5</sup>

(7) a) Syntax:	F <sub>1</sub> >	gia	>	F <sub>2</sub>
	+Neg P			+Neg P
	+ P presupposed			
b) Lexicon:	Piedmontese			Valdotain
	<b>pa</b> :+ Neg P			<b>pa</b> :+Neg P
	+ P presupposed			+/- P presupposed
	<b>nen</b> : + Neg P			

Consider now the second example of section 2, concerning the Italian adverb *presto*, which corresponds to both English *soon* and *early*. The relevant examples, (2) and (3), are repeated here as (8) and (9):

- (8) a) *He will soon have to get up.*  
 b) *He will have to get up early.*

4 It is not clear whether the two *pa* can co-occur (Raffaella Zanuttini, p.c.).

5 This requires not extending to such cases of underspecification the Aspects proposal that “each lexical entry automatically, by convention, contains the feature [-A] for every lexical category A, unless it is explicitly provided by the feature [+A]” (Chomsky 1965,111). The notion of ‘underspecification’ of syntactic features discussed here is different from the phonological and (one type of) morphological notion of ‘underspecification’ discussed in the literature; namely, that concerning those features (like the aspiration of onset stop consonants in English) that are predictable and thus can be expunged from the lexical representation and added through a rule (cf., among others, Archangeli 1984, Farkas 1990, and Steriade 1995). In the cases discussed here, the underspecified features are crucially not added (specified) at all, whether by rule or otherwise. It does bear some similarity, however, with the notion of underspecification employed in Distributed Morphology to account for cases of syncretism. According to Halle’s (1997) Subset Principle “the phonological exponent of a Vocabulary Item is inserted into a position if the item matches all or a subset of the features specified in that position.” (also see Embick and Noyer 2007, §2.4).

- (9) a) *Presto si dovrà alzare.*  
 ‘He will soon have to get up.’  
 b) *Si dovrà alzare presto.*  
 ‘He will have to get up early.’

When *presto* precedes the finite verb it is interpreted as ‘soon’. When it follows the verb it is interpreted as ‘early’.<sup>6</sup> The two *presto* can co-occur:

- (10) *Presto si dovrà alzare presto.* ‘He will soon have to get up early.’

In Cinque (1999) I had suggested that *presto*, qua ‘soon’, encodes “the fact that an event is going to take place *a short while after* some reference time” (p. 97) (cf. *Si rese/ renderà conto presto che lo stavano/stanno imbrogliando* ‘He soon realized/will soon realize that they were/are cheating him’). *Presto*, qua ‘early’, appears instead to be paraphrasable as ‘*a short time after* the beginning of a scale of waking-up (more generally: V-ing) times’ (cf. *Si è alzato presto* ‘He woke up early’). The shared core-component/feature of the two functional projections thus appears to be **a short time after x**. If the lexical specification of *presto* is [**a short time after x**] (x left unspecified), then *presto* will be able to match the two distinct functional projections: the one associated with [**in a short time after x, x a reference time**] and the other associated with [**in a short time after x, x the beginning of a scale of V-ing times**].<sup>7</sup>

Consider next the examples from the substantive lexicon mentioned in section 3.

## 4.2 Substantive Lexicon

The items of the substantive lexicon have components/features that, differently from those of the functional lexicon, do not match components/features of functional heads. Their components/ features rather appear to match the categories with which we interpret/represent the world, broadly taken.

Consider the Italian – Bulgarian contrast shown in Example 1 of section 3. While Italian has two separate lexical items for ‘arm’ and ‘hand’ (*braccio* and *mano*, respectively), Bulgarian has a single lexical item, *raka*, which can refer to either ‘arm’ or ‘hand’. Similarly, while Italian has two separate lexical items for ‘leg’ and ‘foot’ (*gamba* and *piede*, respectively), Bulgarian has just one lexical item, *krak*, which can refer to either ‘leg’ or ‘foot’. I take this to suggest that Bulgarian expresses just the shared component/feature of ‘arm’ and ‘hand’ (namely, ‘upper limb’), and ‘leg’ and ‘foot’ (namely, ‘lower limb’), leaving unspecified what further differentiates ‘arm’ from ‘hand’ and ‘leg’ from ‘foot’. The separate lexical items of Italian for ‘arm’ and ‘hand’ and ‘leg’ and ‘foot’, on the other hand, in addition to specifying the shared component/

6 In English, *early* also has to follow the verb:  
 (i) He < \*early > got up < early >

7 Richard Kayne has suggested a similar analysis in class lectures, also proposing that the differentiating components/features are represented silently.

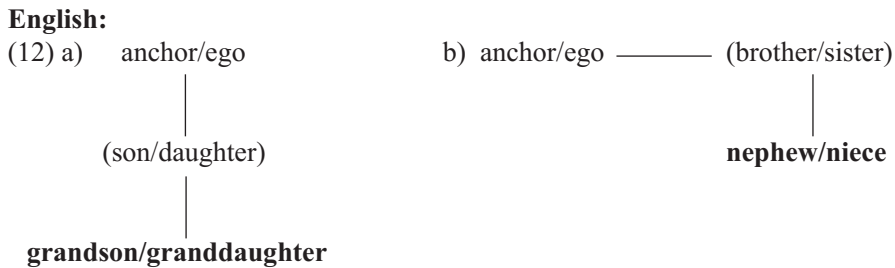


feature, also specify what differentiates ‘arm’ from ‘hand’ and ‘leg’ from ‘foot’. The lexical specifications of the different lexical items of the two languages can thus be represented in first approximation as in (11):

- (11) a) **Italian:** *braccio* ‘arm’ (+upper limb, - extremity)  
*mano* ‘hand’ (+upper limb, +extremity)  
*gamba* ‘leg’ (+lower limb, - extremity)  
*piede* ‘foot’ (+lower limb, +extremity)

- b) **Bulgarian:** *raka* ‘arm’ or ‘hand’ (+upper limb)  
*krak* ‘leg’ or ‘foot’ (+lower limb)

Consider now the second example of section 3. In Italian, a single lexical item, *nipote*, corresponds to English *grandson*, *granddaughter*, *nephew* and *niece*; abstracting away from the male/female distinction (also present in Italian in the determiners that precede the noun: *un/il* (masc.) *nipote*, *una/la* (fem.) *nipote*), *nipote* apparently corresponds in English to two distinct kinship relations, which can be represented as in (12):



- 1<sup>st</sup> line: + descending, - ascending      1<sup>st</sup> line: - descending, - ascending (=horizontal)  
 2<sup>nd</sup> line: + descending, - ascending      2<sup>nd</sup> line: + descending, - ascending

These two kinship relations have, nonetheless, something in common. A degree 2 distance from the anchor/ego. The relation can be made identical if one suspends the directionality of the first line. By leaving unspecified its “descending” component/feature, one can collapse the two kinship relations into one, as in (13), which is precisely what Italian seems to do.<sup>8</sup>

8 If one takes the +descending value of the first line, one gets the ‘grandson/granddaughter’ meaning; if one takes the -descending value, one takes the ‘nephew/niece’ meaning.

It is tempting to take such under-specification of components/features as a way of capturing the cross-linguistic typology of kinship terms. To mention just one example, in Western Dani (Papuan, Trans-New Guinea – Barclay 2008, 61), the lexical word *ombo* means both ‘grandparent’ and ‘grandchild’. In English *grandparent* and *grandchild* have two degrees of distance from the anchor/ego. In the former, both lines are +ascending -descending; in the latter both

**Italian: nipote**

(13) 1<sup>st</sup> line: +/-descending, -ascending

2<sup>nd</sup> line: +descending, -ascending

**5. UNDERSPECIFICATION VS. SILENT ELEMENTS**

In taking an ‘underspecification’ parametric approach to cross-linguistic differences in the lexicons of languages care should be taken to distinguish cases amenable to it from cases arguably involving the presence of silent elements (in one language but not in another), as in Richard Kayne’s recent work. Consider another difference between Italian and Bulgarian, which at first sight appears to be of the same ilk as the preceding ones. While Italian has one word, *molto*, for ‘(very) much’ and another word, *troppo*, for ‘too much’, Bulgarian has a single word, *mnogo*, for both. See, for example the contrast between (14), (15) and (16):

**Italian:**

(14) a) *Non ho bevuto molto.*

‘I didn’t drink much.’

b) *Ha molti libri.*

‘(S)he has got many books.’

c) *Suo figlio è molto stupido.*

‘His son is very (\*much) stupid.’

(15) a) *Ho bevuto troppo.*

‘I drank too much.’

b) *Ha troppi libri.*

‘(S)he has too many books.’

c) *Suo figlio è troppo/\*molto stupido per fare una cosa del genere.*

‘His son is too/very stupid to do such a thing.’

**Bulgarian:**

(16) a) *Toj pie mnogo.*

‘He drinks very much or too much.’

b) *Toj ima mnogo knigi.*

‘He has many or too many books.’

c) *Sinăt mu e mnogo glupav.*

‘His son is very or too stupid.’

---

are -ascending +descending. Western Dani *ombo* thus appears characterizable as underspecified for the +/-ascending, +/-descending components/features (provided that both lines have the same value for such components/features). This line of analysis makes us expect that no single term may cover, say, ‘grandchild’ and ‘cousin’, or ‘nephew/niece’ and ‘cousin’, or ‘grandchild’, ‘nephew/niece’ and ‘cousin’ (‘cousin’ being 3 degrees of distance: 1) + ascending, - descending; 2) - ascending, - descending; 3) - ascending, + descending).

c') *Sināt mu e mnogo glupav za da razbere tova.*  
 'His son is too stupid to understand that.' (another option is to use *tvārde* 'too (much)': *Sināt mu e tvārde glupav za da razbere tova.*)

In this case, thinking of Kayne (2005b, §3.5, 2007), there is reason to believe that the apparent ambiguity of *mного* 'very (much/many)/too (much/many)' is due to the presence of either one of two different silent degree words (*strašno* 'very' and *tvārde* 'too'), as these are the only degree words which are optional in the paradigms (17) and (18), and the only two which are in complementary distribution with *mного* in the paradigm in (19). Consider the following paradigms (Iliyana Krapova, p.c.):

- (17) a) *\*(kolko) mnogo* 'how much' (or simply *kolko*)  
 b) *\*(tolkova) mnogo* 'so much'  
 c) *\*(pò) mnogo* 'more' (or *poveče*)  
 d) *\*(naj) mnogo* 'most'  
 e) *(strašno) mnogo* 'very much/many'  
 f) *(tvārde) mnogo* 'too much/many'
- (18) a) *\*(kolko) mnogo knigi* 'how many books' (or simply *kolko knigi*)  
 b) *\*(tolkova) mnogo knigi* 'so many books'  
 c) *\*(pò) mnogo knigi* 'more books'  
 d) *\*(naj) mnogo knigi* 'most books'  
 e) *(strašno) mnogo knigi* 'very many books'  
 f) *(tvārde) mnogo knigi* 'too many books'
- (19) a) *kolko* (*\*mного*) *glupav* 'how stupid'  
 b) *tolkova* (*\*mного*) *glupav* 'so stupid'  
 c) *pò* (*\*mного*) *glupav* 'more stupid'  
 d) *naj* (*\*mного*) *glupav* 'most stupid'  
 e) *strašno* (*\*mного*) *glupav* 'very stupid'<sup>9</sup> **or** *mного glupav*<sup>10</sup>  
 f) *tvārde* (*\*mного*) *glupav* 'too stupid'<sup>11</sup> **or** *mного glupav*<sup>12</sup>

Thus the ambiguity of (16a-c) is plausibly to be attributed to the presence of a silent degree word; either *strašno* 'very' or *tvārde* 'too' (which cannot be overtly realized within an AP, if *mного* is). Here *mного* is not lexically underspecified. It acquires its apparent ambiguity as a consequence of the independent property of *strašno* 'very' and *tvārde* 'too' to be unpronounced.

9 I.e., *strašno MNOGO glupav* 'very stupid'. Capitals indicate non-pronounced elements.

10 I.e., *STRAŠNO mnogo glupav* 'very stupid'.

11 I.e., *tvārde MNOGO glupav* 'too stupid'.

12 *TVĀRDE mnogo glupav* 'too stupid'.

## 6. CONCLUSIONS

Returning now to the question posed at the beginning (whether and how the cases that we have examined so far can be reduced to one of the parametric actions seen in (1) above), it appears that while the contrast between Italian *molto/troppo* vs. Bulgarian *mnogo* is indeed amenable to the action in (1c) (pronunciation vs. non-pronunciation), the other cases examined in sections 2, 3, and 4 must be attributed to an additional parametric action: underspecification of features in the (substantive and functional) lexicon.

## References

- ARCHANGELI, Diana Bennett (1984) *Underspecification in Yawelmani phonology and morphology*. Ph.D. Dissertation. MIT.
- BARCLAY, Peter (2008) *A grammar of Western Dani*. München: Lincom Europa.
- BORER, Hagit (1984) *Parametric syntax: case studies in Semitic and Romance languages*. Dordrecht: Foris.
- CHOMSKY, Noam (1965) *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.
- CHOMSKY, Noam (1995) *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- CINQUE, Guglielmo (1976) "Mica." *Annali della Facoltà di Lettere e Filosofia dell'Università di Padova* 1, 101–112. [Reprinted in: G. Cinque (1991) *Teoria linguistica e sintassi italiana*. Bologna: Il Mulino.]
- CINQUE, Guglielmo (1999) *Adverbs and Functional Heads. A Cross-linguistic Perspective*. New York: Oxford University Press.
- EMBICK, David/Rolf NOYER (2007) "Distributed Morphology and the Syntax/Morphology Interface." In: G. Ramchand/C. Reiss (eds), *The Oxford Handbook of Linguistic Interfaces*. New York: Oxford University Press, 289–324.
- FARKAS, Donka (1990) "Two Cases of Underspecification in Morphology." *Linguistic Inquiry* 21, 539–550.
- HALLE, Morris (1997) "Distributed Morphology: Impoverishment and Fission." *MIT Working Papers in Linguistics* 30, 425–449.
- KAYNE, Richard (2005a) "Some Notes on Comparative Syntax, with Special Reference to English and French." In: G. Cinque/R. Kayne (eds), *The Oxford Handbook of Comparative Syntax*, 3–69. New York: Oxford University Press.
- KAYNE, Richard (2005b) "A Note on the Syntax of Quantity in English." In: R. Kayne (2005) *Movement and Silence*, 176–214. New York: Oxford University Press.
- KAYNE, Richard (2007) "Several, few and many." *Lingua* 117, 832–858 (also in: R. Kayne (2010) *Comparisons and Contrasts*, 29–56. New York: Oxford University Press).
- RIZZI, Luigi (2009) "Some elements of syntactic computation." In D. Bickerton/E. Szathmáry (eds), *Biological Foundations and Origin of Syntax*, 63–88. Cambridge, Mass.: MIT Press.

- RIZZI, Luigi (2011) "The Elements of Variation: format, locus and acquisition of parametric properties." Talk given at the *50 Years of Linguistics at MIT*, 9 Dec. 2011.
- STERIADE, Donca (1995) "Underspecification and Markedness." In: J. Goldsmith (ed.), *The Handbook of Phonological Theory*, 114–174. Oxford: Blackwell.
- ZANUTTINI, Raffaella (1997) *Negation and Clausal Structure. A Comparative Study of Romance Languages*. New York: Oxford University Press.

### Summary

#### A SOURCE OF PARAMETRIC VARIATION IN THE LEXICON

An influential conjecture concerning parameters is that they can possibly be "restricted to formal features of functional categories" (Chomsky 1995: 6). In Rizzi (2009, 2011) such features are understood as instructions triggering one of the following syntactic actions: (1) External Merge; (2) Internal Merge (Move); (3) Pronunciation/Non pronunciation (the latter arguably dependent on Internal Merge – Kayne 2005a, b). In this article I consider a particular source of parametric variation across languages in the domain of the lexicon (both functional and substantive) which appears to be due to the possibility of underspecifying certain features in some languages. The paradigmatic variation can be characterized as follows: language A has two (or more) lexical items which correspond to just one lexical item in language B.

**Keywords:** parameters, lexicon, underspecification, features

### Povzetek

#### VIR PARAMETRIČNE RAZNOLIKOSTI V LEKSIKONU

Ena izmed vplivnih domnev o naravi parametrov predvideva, da so parametri »omejeni zgolj na formalne oznake funkcijskih kategorij« (Chomsky 1995: 6). Rizzi (2009, 2011) predlaga, da te oznake služijo kot navodila jezikovni izgradnji, ki sprožijo eno od naslednjih skladijskih operacij: (1) Sestavi (External Merge); (2) Premakni (Internal Merge, Move); (3) Izgovorjava/Ne-izgovorjava (Pronunciation/Non pronunciation), pri čemer je zadnja operacija verjetno odvisna od operacije Premakni (Kayne 2005a, b). Članek obravnava specifičen vir parametrične raznolikosti v domeni leksikona (tako funkcijskega kot leksikalnega), ki je, kot kaže, posledica možnosti podspecifikacije določenih oznak v nekaterih jezikih, in pokaže, da paradigmatično variacijo lahko opišemo na sledeč način: jezik A ima dva (ali več) leksikalnih elementov, ki ustrezajo samo enemu elementu v jeziku B.

**Ključne besede:** parametri, leksikon, podspecifikacija, oznake



## EXPERIMENTAL EVIDENCE FOR NEG-RAISING IN SLAVIC\*\*\*

### 1 INTRODUCTION

Neg-raising (NR) is an interpretational phenomenon: in biclausal structure like (1), negation of the root predicate (*think*) is most saliently interpreted on the embedded predicate, i.e., (1a) is normally interpreted as meaning (1b). Not every predicate demonstrates this property, however. For instance, verbs of communication are not neg-raisers, cf. the intuition that *John says that it's not raining* does not follow from *John doesn't say that it's raining*.

- (1) a) *John doesn't think that it's raining.*  
b)  $\sim$  *John thinks that it's not raining.*

The NR interpretation is not predicted by the standard semantics of propositional attitudes (see, e.g., Hintikka 1969). For example, the verb *think* is formalised as a universal quantifier over possible worlds, restricted to some modal base – see (2). This predicts that (2a) does not entail (2b). Thus, it seems that the standard semantics correctly characterise communication verbs, but they fail for verbs like *think*. To put it differently: if we want to formalise the reasoning from (1a) to (1b), something more must be said about NR predicates.

- (2)  $\|think\|(p)(a)(w) = \forall w' \in M(w, a)[p(w')]$   
a)  $\neg [\forall w' \in M(w, a)[p(w')]]$   
b)  $\rightarrow [\forall w' \in M(w, a)[\neg p(w')]]$

While *think* is probably the most well-known example of NR predicates, other propositional attitude (PA) verbs are neg-raisers, too. Since Horn (1989), it has been

---

\* docekal@phil.muni.cz

\*\* j.dotlacil@gmail.com

\*\*\* We would like to thank the audiences at SinFonIJA 8 and FDSL 11, the organizers, editors and especially two anonymous reviewers of *Linguistica* for their useful comments and questions. We would further like to thank various people who commented on previous versions of this paper, especially Viola Schmitt, Jacek Witkoś and Joanna Błaszczak. Any flaws remain our responsibility.

common to split NR predicates into five classes, and we will use this categorisation in the current article as well:

- a) intention (*want, intend, ...*),
- b) obligation (*advise, should, ...*),
- c) perception (*seem, appear, ...*),
- d) opinion (*know, believe, ...*),
- e) probability (*probable, likely, ...*).

Most research studying NR focuses on English.<sup>1</sup> One notable exception is Bošković & Gajewski (2009) (henceforth: B&G) and Dočekal (2014), which consider Slavic languages without articles. B&G argue that NR is absent in these languages and demonstrate how the finding follows from Gajewski's (2005) NR theory, which is a modern, sophisticated, and updated version of Bartsch's (1973) presuppositional approach to NR. In contrast to B&G, Dočekal (2014) argues that predicates of intention and obligation (a and b above) pass standard tests for neg-raisers in Czech (a Slavic language lacking articles), that is, Czech has NR predicates, at least in Dočekal's introspection. There are three tests Dočekal (2014) considers.

As discussed already by Lakoff (1969) (see also Gajewski 2005, 2011 for a recent discussion), strict negative polarity items (NPIs) can be licensed by negated NR predicates. We will discuss properties of strict NPIs later in the paper. At this point, it suffices to say that *until*, modifying punctual events, is standardly considered an example of a strict NPI and that we see that *until tomorrow* is possible when embedded under negated NRs, (3a), and ungrammatical under non-NR predicates, see (3b).

- (3) a) *Bill didn't think that Mary would leave until tomorrow.*
- b) *\*Bill didn't say that Mary would leave until tomorrow.*

The second standard test for NR-hood is the cyclicity of NR inferences. For example, (4a) is most saliently understood as (4b), in which the negation is interpreted on the most embedded predicate. This inference is broken if any of the predicates is a non-NR predicate.

- (4) a) *I don't believe that he wants me to think that he wrote it.*
- b) *I believe that he wants me to think that he didn't write it.*

The third standard test for NR-hood concerns the inference that NR predicates yield when appearing in the scope of a negated universal, e.g. (5a) implies (5b).

- (5) a) *Not every student thinks that John is a good teacher.*
- b) *There are some students who think that John isn't a good teacher.*

---

1 For recent theories of NR, see Gajewski (2007), Romoli (2013), and Collins and Postal (2014).

According to Dočekal’s intuitions, Czech NR predicates pass all three aforementioned tests. However, judgments related to NR are subtle and difficult. For this reason, we want to consider the existence of NR in Czech experimentally. We argue that the experiment confirms Dočekal’s (2014) position and provides evidence against the empirical claim of Bošković & Gajewski (2009). We interpret our data in the scalar framework of NR (originally coined by Horn 1989 and formalised currently in Romoli 2012, 2013).

## 2 EXPERIMENT

### 2.1 Method

#### 2.1.1 Introduction

The experiment testing NR predicates in Czech consisted of two parts:

- a) an acceptability judgment task,
- b) an inference task.

#### 2.1.2 Acceptability Task: Materials

The acceptability task tested how Czech native speakers accept so-called strict negative polarity items, or NPIs (Zwarts 1998). Strict NPIs are licensed in the subset of environments that license weak NPIs (on the latter, see, e.g., Ladusaw 1979). What is important for us is that it has been demonstrated that strict NPIs can be licensed by clausemate negation or by negated NR predicates, and consequently, they have often been used as tests for NR-hood (see, e.g., Gajewski, 2003). We will consider two types of strict NPIs:

- a) *ani jeden* ‘even one’ + N type (like *ani jedna ovce* ‘even one sheep’),
- b) *až do* ‘until’ + time expression

The acceptability of strict NPIs was judged on a Likert scale from 1 (*absolutně nepřijatelná věta*, ‘sentence completely unacceptable’) to 5 (*věta je naprosto v pořádku*, ‘sentence fully acceptable’).

We tested the acceptability of strict NPIs in five conditions:

- (A) in positive sentences,
- (B) in simple negative sentences,
- (C) in clauses embedded under negated NR predicates of intention (e.g., *want*) and judgment/obligation (e.g., *advise*),
- (D) in clauses embedded under negated NR predicates of opinion (e.g., *believe*),
- (E) in clauses embedded under negated non-NR predicates (prototypically verbs of communication and causation – *say, make, ...*)

We split NR predicates into two conditions, C and D, to see whether or not different semantic classes of NR predicates behave differently. This is of interest since B&G tested NR across Slavic languages only on a representative of the opinion class, and it is



conceivable that their conclusion were correct for that class, while group C does show properties of NR in Czech.

There were 40 items in the acceptability task. Five lists were constructed out of the items and conditions in such a way that each item appeared only once in any list and the five conditions were ‘cycled’ through the lists (repeated Latin-square design). Half of the items were tested with the strict NPI *ani jeden*, while the other half used the strict NPI *až do*. All of the items in the acceptability task were tested in all five conditions (A) – (E) from above. One example of an item with strict the NPI *ani jeden* is presented in (6) (verbs and strict NPIs are boldfaced in the example, but they were not marked in the experiment).

- (6) a) **Ztratila se ani jedna ovce.**  
 lost SE even one sheep  
 ‘A single sheep is missing.’
- b) **Neztratila se ani jedna ovce.**  
 neg-lost SE even one sheep  
 ‘Not a single sheep is missing.’
- c) **Nový bača v Tatrách nechce, aby se ztratila ani jedna ovce.**  
 new shepherd in Tatra neg-wants C-SUBJ SE lost even one sheep.  
 ‘The new shepherd in the Tatra mountains does not want a single sheep to be missing.’
- d) **Nový bača v Tatrách si nemyslí, že se ztratila ani jedna ovce.**  
 new shepherd in Tatra SI neg-think C-IND SE lost even one sheep  
 ‘The new shepherd in the Tatra mountains does not think that a single sheep is missing.’
- e) **Nový bača v Tatrách neříká, že-IND se ztratila ani jedna ovce.**  
 new shepherd in Tatra neg-say C SE lost even one sheep  
 ‘The new shepherd in the Tatra mountains does not say that a single sheep is missing.’

Since we deal with strict NPIs, we expect that negative sentences, like (6b), are more acceptable than positive sentences, like (6a). However, the most interesting case for us is the contrast between (6c) and (6d), on the one hand, and (6e), on the other hand. If Czech has NR predicates, we expect that the examples (6c) and (6d), in which strict NPIs are licensed, are more acceptable than (6e), in which the licensing of strict NPIs should not take place.

Careful readers might have noticed that embedded clauses in Conditions C and D in the example above differ in their mood: condition C uses subjunctive (glossed as SUBJ and realised on the complementiser), while Condition D and Condition E use indicative (glossed as IND on the complementiser). In the experiment, not every item had this distribution of mood. The items were constructed in such a way that (i) subjunctive mood was always used in Condition C because NR predicates of that group disallow any other mood, and (ii) 10 items in Condition D had an indicative mood, while the

remaining 30 items used subjunctive. Furthermore, Condition E (non-NR predicates) had 15 items in the subjunctive mood and the remaining 25 items in the indicative mood. This arbitrary division of subjunctive and indicative was caused because we did not consider mood a relevant factor when designing the experiment. We will return to this issue in section 3.3.

The verbs we used in testing the (C) environment were intention/judgment/obligation verbs like the following: *chtít* ‘want’, *hodlat*, *dopustit* ‘allow’, *mít v úmyslu* ‘have an intention’, *přát si* ‘wish’, *vyžadovat* ‘require’, *potřebovat* ‘need’, *usilovat* ‘strive’, *radit* ‘advise’, *doporučovat* ‘recommend’, and *navrhovat* ‘propose’. Examples of verbs from the (D) environment are opinion verbs like: *myslet* ‘think’, *věřit* ‘believe’, *předpokládat* ‘suppose’, *představovat* ‘imagine’, *očekávat* ‘expect’, *uvažovat o* ‘speculate’, *domnívat se* ‘assume’, *soudit* ‘judge’, and *spoléhat se* ‘count on’. Lastly, in the (E) environment, we used non-NR predicates such as *říkat* ‘say’, *slyšet* ‘hear’, *tvrdit* ‘assert’, *rozhlášovat* ‘rumour’, *naznačit* ‘indicate’, *prozradit* ‘reveal’, *sdělit* ‘tell’, *zavolat* ‘call’, *napsat* ‘write’, *způsobit* ‘cause’, *vyrozumět* ‘inform’, *nutit* ‘force’, *číst* ‘read’, and *chápat* ‘understand’.

There were also 30 fillers in the acceptability task.

All of the stimuli were presented to each participant in a random order.

### 2.1.3 Inference Task: Materials

In the inference task, which always followed the acceptability task, we tested whether or not Czech native speakers found three entailments valid. The entailments are usually taken as another test for NR-hood (see Horn 1989 a.o.). The first condition tested the validity of NR itself ( $\neg \text{NR}[P] \rightsquigarrow \text{NR}[\neg P]$ ) illustrated in (7). Participants were asked whether or not (7b) follows from (7a). Note that the sentence is similar to the one used in the acceptability task, but there is no strict NPI used here. The second condition tested whether or not participants accept cyclic neg-raising ( $\neg \text{NR}_1[\text{NR}_2[P]] \rightsquigarrow \text{NR}_1[\text{NR}_2[\neg P]]$ ). For example, they were asked whether or not (8b) follow from (8a). The final condition targeted the existential wide scope inference ( $\neg \forall x \text{NR}_1[\text{NR}_2[P]] \rightsquigarrow \exists x \text{NR}_1[\text{NR}_2[\neg P]]$ ), i.e., whether or not (9b) follows from (9a).

- (7) a) *Nový bača v Tatrách nechce, aby se ztratila jediná ovce.*  
 new shepherd in Tatra neg-wants C-SUBJ SE lost single sheep  
 ‘The new shepherd in the Tatra mountains does not want a single sheep to be missing.’  
 b) *Nový bača v Tatrách chce, aby se neztratila jediná ovce.*  
 new shepherd in Tatra wants C-SUBJ SE neg-lost single sheep  
 ‘The new shepherd in the Tatra mountains wants no sheep to be missing.’
- (8) a) *Myslivci nevěří, že nový bača v Tatrách chce, aby se ztratila jediná ovce.*  
 hunters not-believe that new shepherd in Tatra wants C-SUBJ SE lost single sheep  
 ‘The hunters do not believe that the new shepherd in the Tatra mountains wants a single sheep to be missing.’

- b) *Myslivci věří, že nový bača v Tatrách chce, aby se neztratila jediná ovce.*  
 hunters believe that new shepherd in Tatra wants C-SUBJ SE neg-lost single sheep  
 ‘The hunters believe that the new shepherd in the Tatra mountains wants no sheep to be missing.’
- (9) a) *Ne všichni myslivci věří, že nový bača v Tatrách chce, aby se ovce měly dobře.*  
 not all hunters believe that new shepherd in Tatra wants C-SUBJ SE sheep had well  
 ‘Not all the hunters believe that the new shepherd in the Tatra mountains wants sheep to prosper.’
- b) *Někteří myslivci věří, že nový bača v Tatrách chce, aby se ovce neměly dobře.*  
 some hunters believe that new shepherd in Tatra wants C-SUBJ SE sheep neg-be good  
 Some hunters believe that the new shepherd in the Tatra mountains wants sheep not to prosper.’

Twenty items were used in the inference task. The items were split into three lists, so that each item appeared only once in each list and the conditions were ‘cycled’ through the items (in the same manner as in the acceptability task). Apart from the experimental items, the inference task also included 30 fillers. The experimental fillers were uncontroversial cases of either deductively valid inferences (like *modus ponens*: paraphrase of premises – *If it rains, the streets are wet; It is raining; conclusion: The streets are wet*) and cases of deductively non-valid inferences (logical fallacies). It is unlikely that the fillers had any impact on the inferences in the items and vice versa. The stimuli were presented in a random order.

#### 2.1.4 Procedure and Participants

The experiment was prepared in Ibsen and participants engaged with the experiment online. The experiment began with instructions, and these were followed by practice items and the acceptability task. Afterwards, the inference task was presented. Sixty native speakers of Czech, mostly students of linguistics in Brno, participated in the experiment, and they received course credit for their participation.

#### 2.2 Results of Acceptability Task

The fillers in the acceptability task were uncontroversially grammatical or ungrammatical according to our intuitions, and we used them to check that participants understood the task. More concretely, we checked whether or not the average of each participant’s responses to ungrammatical fillers was lower than the average of his or her responses to grammatical fillers. Indeed, every participant passed this test (there was at least a one point difference between the two averages). We kept all the participants for the subsequent analysis.

Responses in the acceptability task were modeled by mixed-effects ordered probit regression in the R package `ORDINAL`.<sup>2</sup> The model had one predictor, Condition

2 The advantage of using ordered probit regressions as compared to the more familiar linear regression models is that the former only assumes that responses 1-5 are ordered, but does not

(condition C, i.e., *want* type of neg-raisers, was the reference level). The model also included subject and item slope+intercept random effects. The statistical outcome was the following:<sup>3</sup>

- a) simple negated sentences (Condition B) with strict NPIs were judged as better than NR predicates from the reference level ( $\beta = 3.2, z = 7.3, p < .001$ ),
- b) simple positive sentences (Condition A) with strict NPIs were judged as worse than any NRs ( $\beta = -1.5, z = -9.2, p < .001$ ),
- c) sentences with negated non-NR predicates (condition E) were worse than NR predicates ( $\beta = -0.8, z = -5.6, p < .001$ )
- d) there was no significant difference between two types of NR predicates (Condition C and Condition D) ( $p > .1$ )

The boxplots of the acceptability ratings depending on the five conditions are charted in Figure 1 and we think this graphical summary visually matches what has been found using inferential statistics. We conclude that the statistical analysis is compatible with the following hypotheses about Czech:

- a) *ani jeden* ‘even one’ and *až do* ‘until’ are strict NPIs (if they were weak NPIs, they would be grammatical in Condition E, negated non-NR predicates; if they were not polarity items, we would expect that they were acceptable in Condition A, positive sentences; both predictions are incompatible with our findings),
- b) Czech has a class of NR verbs – NR verbs of intention, obligation, and opinion are able (with some interesting variation, discussed further) to license strict NPIs in their embedded clauses, unlike Czech non-NR predicates, hence the difference between Condition C and D, on the one hand, and Condition E, on the other hand. At the same time, there is a difference between NR predicates and clause-mate negation in that strict NPIs are less acceptable in the former environment (cf. the difference between Condition C and D, on the one hand, and Condition B, on the other hand).

---

presuppose that the points in the scale are evenly spaced. Rather, the distance between any two points is modeled from data. This is good, since it is possible that participants do not treat Likert scales as linear, e.g., the difference between 1 and 2 may not be the same as the difference between 2 and 3.

3 When interpreting the outcome, it should help to know that thresholds between responses had the following estimates: 1|2:  $\beta = -1.1$ ; 2|3:  $\beta = -0.06$ ; 3|4:  $\beta = 0.7$ ; 4|5:  $\beta = 1.6$ .

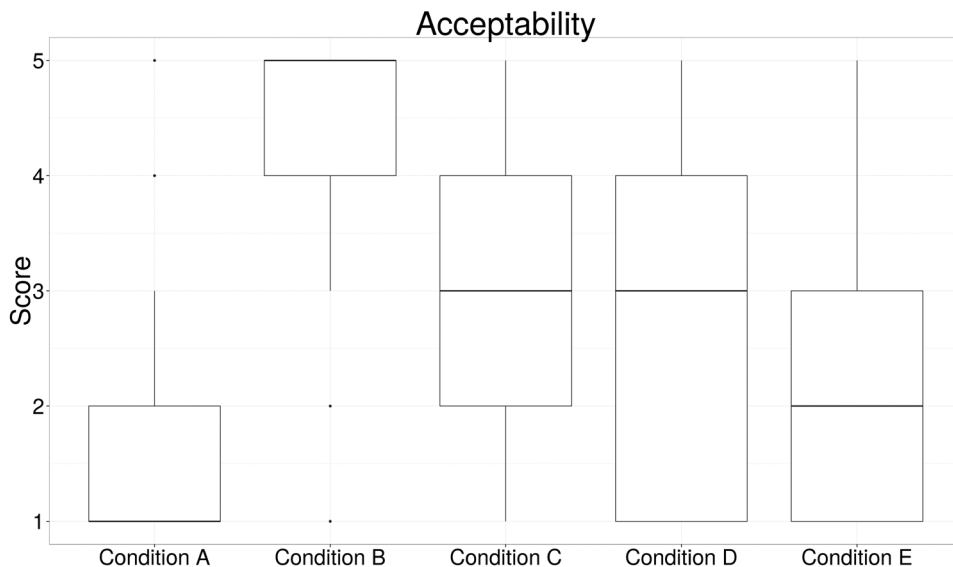


Figure 1: Results of the Acceptability Task

Two outstanding issues were investigated post-hoc. First, as noted above, we used two types of strict NPIs in the experiment. We wanted to see to which extent the two NPIs differ in acceptability. Secondly, two moods were used in NR predicates (subjunctive and indicative). Any differences between them in licensing strict NPIs were investigated. Importantly, the experiment was not designed to address either of the issues, so these results must be taken only as preliminary.

To study the first issue, we added a new factor to the model, NPI-type (with two factors, *ani jeden* ‘even one’ and *až do* ‘until’ the former was the reference level) and its interaction with Condition (as before, Condition C was the reference level). The model had the same random-effect structure as the previous one. What we see is that in Condition C, ‘until’ is more acceptable than ‘even one’ ( $\beta = 0.4$ ,  $z = 2.4$ ,  $p < .05$ ). As before, positive sentences and non-NR predicates are judged as worse ( $\beta = -1.7$ ,  $z = -8.3$ ,  $p < .001$  and  $\beta = -1.1$ ,  $z = -5.7$ ,  $p < .001$ , respectively), and negative sentences as better ( $\beta = 4.7$ ,  $z = 10.4$ ,  $p < .001$ ). Condition D was judged as slightly worse with ‘even one’ ( $\beta = -0.4$ ,  $z = 2.5$ ,  $p < .05$ ). Importantly, NPI type interacted with Conditions: ‘until’  $\times$  Condition D and ‘until’  $\times$  Condition E led to higher acceptance ( $\beta = 0.6$ ,  $z = 2.6$ ,  $p < .01$  and  $\beta = 0.6$ ,  $z = 2.3$ ,  $p < .05$ , respectively), while ‘until’  $\times$  Condition B (simple negative sentences) was less acceptable ( $\beta = -3.2$ ,  $z = -6.4$ ,  $p < .001$ ). The findings indicate that ‘even one’ is a better representative of strict NPIs than ‘until’ is. This is because ‘even one’ is fully acceptable with clausemate negation and it is somewhat degraded with negated NR predicates and even more so with negated non-NR

predicates, which is a behaviour one would expect from strict NPIs (disregarding the somewhat degraded status with negated NR predicates, which will be discussed later). The situation is more complex for ‘until’. As was the case for ‘even one’, the best environment for ‘until’ is clausemate negation, Condition B. However, compared to ‘even one’, ‘until’ is *significantly less* acceptable there, which is surprising if it was simply a plain, strict NPI. Even more interestingly, ‘until’ is *more* acceptable with negated NR and non-NR predicates (Conditions C, D and E) than ‘even one’ is. To take a step back, it seems that ‘until’ lacks an environment that would fully license it, unlike ‘even one’, at least as far as our experiment is concerned. This casts doubts on its status as a strict NPI and on using it as a testbed for NR-hood of predicates. Importantly for us, even when we restrict our attention only to ‘even one’-NPIs, we still see that NR predicates (Conditions C and D) are better licensors than non-NR predicates or positive sentences.

The second post-hoc issue concerned the subjunctive/indicative difference. To study this factor, we analysed only Conditions that used both moods (Condition D, NR predicates of opinion, and Condition E, non-NR predicates). Analysing the data in a mixed-effects ordered probit model with two fixed effects, Condition and Mood, we found that the indicative mood degraded acceptability ( $\beta = -1.0$ ,  $z = -4.7$ ,  $p < .001$ ), and so did non-NR predicates ( $\beta = -0.8$ ,  $z = -5.4$ ,  $p < .001$ ). The interaction of the two factors was not significant ( $p > .1$ ). To sum up, we see that mood plays a role in licensing strict-NPIs and unfortunately, this factor was not fully considered when we designed the experiment. At the same time, mood clearly cannot be the sole factor at play, since even after we added it to the model, the difference between NR and non-NR predicates remained a highly significant predictor of the acceptability of strict NPIs.

### 2.3 Results of Inference Task

The inference task was analysed in a mixed-effects logistic regression, where 1 = inference follows and 0 = inference does not follow, using R package LME4. Recall that the inference task consisted of three conditions:

Condition I:  $\neg \text{NR}[P] \sim \text{NR}[\neg P]$

Condition II:  $\neg \text{NR}_1[\text{NR}_2[P]] \sim \text{NR}_1[\text{NR}_2[\neg P]]$

Condition III:  $\neg \forall x \text{NR}_1[\text{NR}_2[P]] \sim \exists x \text{NR}_1[\text{NR}_2[\neg P]]$

The model had one fixed factor, Condition. The model also included intercept-only subject and item random effects (more complex models did not converge in LME4). In Condition I (the NR inference), the answer 1 (inference follows) was used in 65% of all cases, which was significantly higher than a chance (prob = 0.5) ( $\beta = 0.9$ ,  $z = 3.3$ ,  $p = .001$ ),<sup>4</sup> so for Condition I, we can safely say that NR reasoning was preferred.

For Condition II (cyclic NR) and Condition III (existential wide scope), response 1 (inference follows) was used in 49% and 48% of all cases respectively, which was not statistically different from chance (prob = 0.5) ( $p > 0.1$ ). These mixed results are

---

4 These values are based on a mixed-effect logistic regression with one factor, Condition, and subject and item intercept-only random effects.

unexpected in all of the previous theoretical accounts of NR, whether they are based on presuppositions or implicatures. We suspect that two factors play a role here:

- a) Conditions II and III are more complex than acceptability task and Condition I. This was particularly clear from response times – Conditions II and III took subjects 2-3 times longer than Condition I. It is reasonable to deduce that subjects became lost in complex sentences.
- b) Conditions II and III always used the indicative mood in the first embedded clause. This will likely impede the NR-type inference, given that the indicative mood also blocks the licensing of strict NPis by NR predicates (see Section 2.2).

### 3. ANALYSIS

We will now consider a framework that can explain all the main findings of our experiment. In particular, there are two asymmetries that our theory of NR will have to address:

- a) NR predicates of intention, obligation, and opinion are better licensors of strict NPis than non-NR predicates.
- b) NR predicates are nevertheless worse strict-NPI licensors than clausemate negation.

We will consider a scalar approach to NR to capture both findings.

#### 3.1 The Scalar Approach to NR

The scalar approach to NR was developed by Romoli (2012, 2013). In Romoli’s approach, NR predicates are equipped with the set of alternatives that consists of the NR predicate itself, as well as the version in which the NR predicate is substituted by a predicate with excluded middle inference. An example is given in (10) for the NR predicate *want* and in (11) for *think*. The member in the set written as the first has the NR predicate itself, while the second member is the predicate with the excluded middle inference. The second member can be paraphrased as being opinionated in the case of *think*, or having a desire in the case of *want*. Note that the (10a) and (11a) versions are just informal descriptions of (10b) and (11b).

- (10) a)  $\text{Alt}(\text{want}(p)(x)) = \{\text{want}(p)(x), \text{have a desire as to whether}(p)(x)\}$
- b)  $\text{Alt}(\text{want}(p)(x)) = \{\Box_x[p], [\Box_x[p] \vee \Box_x[\neg p]]\}$
- (11) a)  $\text{Alt}(\text{think}(p)(x)) = \{\text{think}(p)(x), \text{have an opinion as to whether}(p)(x)\}$
- b)  $\text{Alt}(\text{think}(p)(x)) = \{\Box_x[p], [\Box_x[p] \vee \Box_x[\neg p]]\}$

As an example, consider the computation of the meaning of the sentence *Susan wants to sleep*. We will write the interpretation as in (12), in which  $\Box$  is the translation of the modal verb *want* (all the worlds compatible with Susan’s wishes).

- (12)  $\Box_{\text{susan}}[\text{sleep}(\text{susan})]$

Crucially, this interpretation can be further strengthened. Romoli (2013), following Chierchia (2004), assumes that propositions are strengthened by an exhaustivity operator, EXH, similar in its meaning to *only*. EXH is lexically specified as affirming the proposition to which it attaches and negating excludable alternatives. Excludable alternatives are alternatives that can be negated without contradicting the basic meaning, see (13).

- (13) a)  $EXH(Alt(p))(p)(w) = p(w) \wedge \forall q \in Excl(p, Alt(p))[\neg q(w)]$   
 b)  $Excl(p, Alt(p)) = \{q \in Alt(p): \lambda w[\neg q(w)] \cap p \neq \emptyset\}$

Coming back to our example, we can indicate the basic meaning and the meaning of negated alternatives as in (14). What is worth noting is that neither negated alternative is excludable. Both members of the set contradict the basic meaning. This should be obvious for the first member of the set. The second member can be rewritten as  $\neg \square_{susan}[sleep(susan)] \wedge \square_{susan}[\neg sleep(susan)]$ , which also clearly contradicts the basic meaning. Because of that, alternatives do not play any role in this case and consequently, (14) results as the final meaning of the sentence.

- (14) Basic:  $\square_{susan}[sleep(susan)]$   
 Negated alternatives:  
 $\{\neg \square_{susan}[sleep(susan)], \neg[\square_{susan}[sleep(susan)] \vee \square_{susan}[\neg sleep(susan)]]\}$

The situation changes when we consider the sentence *Susan does not want to sleep*. The negated alternatives are as shown in (15b), where  $p$  is an abbreviation of  $sleep(susan)$ . Now, the second alternative (the disjunction of two propositions) is excludable, since it does not contradict the basic meaning. In fact, it strengthens it, as shown in (15b). The resulting interpretation is that Susan wants not to sleep, which is the neg-raising inference.

- (15) a) Basic meaning:  $\neg \square_{susan}[sleep(susan)]$   
 a) Negated alternatives =  $\{\neg \neg \square_{susan}[p], \neg \neg [\square_{susan}[p] \vee \square_{susan}[\neg p]]\}$   
 $= \{\square_{susan}[p], [\square_{susan}[p] \vee \square_{susan}[\neg p]]\}$   
 b)  $\|EXH\| (\neg want_s[p]) = \neg \square_{susan}[p] \wedge [\square_{susan}[p] \vee \square_{susan}[\neg p]] \models want_s[\neg p]$

The consequence of exhaustification of NR propositions is that negation is interpreted as having low scope. To put it more abstractly,  $\neg NR(p)$  plus the alternative  $NR(p) \vee NR(\neg p)$  entails  $NR(\neg p)$ .

Why should this entailment matter for strict NPIs, however? An answer to this question depends on what we believe the exact mechanism for licensing strict NPIs to be. Currently, three standard approaches are usually considered (see Gajewski 2005, 2011 for summary and details). Here, we will use that of Zwarts (1998): strict NPIs are licensed by anti-additive functions. Anti-additive functions are defined in (16).



- (16) A downward-entailing function  $f$  is anti-additive iff for any  $a$  and  $b$  in the domain of  $f$ ,
- $$f(a) \text{ and } f(b) \leftrightarrow f(a \text{ or } b)$$

Consider how (16) works for negation. (17b) follows from (17a) and vice versa, hence, we can conclude that negation is anti-additive and can license strict NPIs, which is correct. The same conclusion is shown more abstractly using propositional logic in (17c) and (17d). That (17c) and (17d) are equivalent is a straightforward consequence of de Morgan's law.

- (17) a) *It didn't rain and it didn't snow.*  
 b) *It didn't rain or snow.*  
 c)  $\neg p \wedge \neg q$   
 d)  $\neg[p \vee q]$

Crucially, the same conclusion holds for NR predicates. More concretely, (18b) follows from (18a) and vice versa, or, using modal logic, (18d) follows from (18c) and vice versa. To see the latter, notice that (18c) says that in no worlds does  $p$  hold and in no worlds does  $q$  hold. Then, however, it follows that there are no worlds in which  $p \vee q$  holds. Similarly, if there are no worlds in which  $p \vee q$  holds then it must be that in no worlds does  $p$  hold and in no worlds does  $q$  hold.

- (18) a) *Susan does not want to sleep and she does not want to dance.*  
 b) *Susan does not want to sleep or dance.*  
 c)  $\Box\neg p \wedge \Box\neg q$   
 d)  $\Box\neg(p \vee q)$

For the equivalence of (18c) to (18d), we made use of the NR-hood of the predicate *want*. Had we not done so, anti-additivity would not be maintained. In other words, (19b) does not follow from (19a). To see that, consider the following. (19a) is true if there is a world in which  $p$  is not true and a world in which  $q$  is not true. However, (19b) requires that there is a world in which neither  $p$  nor  $q$  is true. That does not follow. In particular, if  $p$  is false only in the world  $w_1$  and  $q$  is false only in the world  $w_2$ , then (19a) is true and yet, (19b) is false.

- (19) a)  $\neg\Box p \wedge \neg\Box q$   
 b)  $\neg\Box(p \vee q)$

If this interpretation of NR and strict NPIs are correct, it follows that the NR inference (and with it, strict NPI licensing) should depend on various factors. First of all, the alternatives must reach the EXH operator. Secondly, the alternatives must be relevant (see Romoli 2013, Sect. 7). Consider (20). The alternatives are relevant if (20) is an answer to a question under discussion (QUD) such as *What does the new shepherd want*

*his sheep to do?* However, the exhaustification might not be an option. We will discuss this possibility in the next section.

(20) *The new shepherd in the Tatra mountains doesn't want even one sheep to be missing.*

The last question we must address is why we see no NR-like inferences with non-NR predicates. Romoli's answer, which we again follow here, is that alternatives triggering NR are absent for non-NR predicates. For example, the non-NR predicate *be certain* has universal and existential quantification over possible worlds as its alternatives. This is shown in (21). Negating the existential alternative of (21) yields the inference demonstrated in (21a). The full meaning is represented in (21b). Applied to the data in our experiment: we observed that negated non-NR predicates are unable to license strict NPIs in embedded clauses. This follows since non-NR predicates do not trigger the low-scope interpretation of negation, even when exhaustified, and the anti-additive inference is not valid.

(21) *John isn't certain that Mary will arrive.*

a)  $\sim$  It's possible for John that Mary will arrive.

b)  $\neg\Box_j[p] \wedge \Diamond_j[p]$

### 3.2 NR Suspension – NR vs. Simple Negated Sentences

We now turn to the question as to why strict NPIs are fully licensed by clausemate negation, while negated NR predicates are worse licensors. Gajewski (2007) observes that, in English, NR inferences can be suspended if the auxiliary in the main clause is stressed. *John DOESN'T think that Fred left* can be used to express that John is not sure about Fred's whereabouts. We believe that similar suspension can take place in Slavic.

One way to capture the suspension of inferences is to say that only relevant alternatives can yield inferences, and what counts as relevant can be modeled using questions under discussion (QUD). In this perspective, every sentence can be seen as an answer to its (implicit or explicit) QUD. QUDs, in turn, are questions that partition the common ground. Relevance is then defined as follows (following Romoli 2013):

(22) Relevance: A proposition  $p$  is relevant to a question  $Q$  iff  $p$  is (contextually equivalent with) the union of some subset of  $Q$ .

An example of a QUD is found in (23). This question creates a partition like (20a): affirmative propositional attitude, negative propositional attitude, and ignorance. In such a context, the alternative triggering low-scope negation is relevant because it is equivalent with  $c_3$ .

(23) *What does the new shepherd want his sheep to do?*

a)  $Q = \{c_1 = \text{want}_s[p], c_2 = \text{want}_s[\neg p], c_3 = \neg(\text{want}_s[p] \vee \text{want}_s[\neg p])\}$

Given that QUDs affect inferences, they can modulate licensing of strict NPIs. Consider (24), in which, according to our intuitions, the stress on the verb blocks the strict NPI in the embedded clause (a proper experimental study that would take intonation into account would be needed here, of course). The focus is the affirmative/negative polarity of the clause. (There is another interpretation, in which the verb itself is the focus, but that is irrelevant for now.) Theoretically, we model this through the QUD in (25), which leads to the partition in (25a). In this case, the alternative triggering low-scope negation, namely,  $\neg(\textit{think}_s[p] \vee \textit{think}_s[\neg p])$ , is not relevant because it is not equivalent to any member in the partition. Consequently, the crucial inference licensing strict NPIs in the embedded clause is not calculated.<sup>5</sup>

(24) *Nový bača si NEMYSLÍ, že se ztratila \*ani jedna ovce.*  
 new shepherd SE DOESN'T-think that SE disappeared even-one one sheep

(25) *Does the new shepherd think that one sheep dissapeared?*  
 a)  $Q = \{c_1 = \textit{think}_s[p], c_2 = \neg \textit{think}_s[p]\}$

This argumentation explains why simple negated sentences (Condition B) are always better as strict NPI licensors: their licensing ability is not dependent on exhaustification and hence is not sensitive to context manipulation. One might wonder, however, why Czech and English differ. In English, licensing strict NPIs by NR predicates has never been questioned, as far as we know. In contrast to this, Slavic NRs have a less clear status, as seen in the disagreement in the literature (Bošković & Gajewski 2009 vs. Dočekal 2014), and as also visible in the lower scores of NR predicates in our acceptability study.

One option is that different morphosyntax of the two markers is to blame. English negation triggers *do*-support and thus, QUD targeting polarity can be straightforwardly marked by stressing the negation itself (plus its host, the semantically vacuous *do*). This is not possible in Czech, in which (the clausal) negation is a bound morpheme, must be attached to the verb, and cannot be independently stressed.<sup>6</sup> Since it is not possible to unambiguously mark the QUD, targeting polarity, Czech speakers are likely to consider this interpretation without any specific signal. Consequently, Czech speakers might suspend the crucial inference more freely.

### 3.3 Subjunctive vs. Indicative

We noted in our discussion of the experiment that two factors affect the licensing of strict NPIs in embedded clauses: one was the type of the matrix predicate (NR or

5 Romoli (2013) uses a different approach for licensing strict NPIs, following Gajewski (2011). Somewhat simplified, the theory states that strict NPIs are licensed only if the meanings strengthened by the application of EXH are downward-entailing. This approach would predict that cases like (24) allow strict NPIs.

6 This is also the reason that the example (21), with the stress on verb, can be understood as focusing the polarity of the sentence or the verb itself.

not-NR), and the analysis of that effect was provided in the previous section. The second factor is the mood in the embedded clause (subjunctive vs. indicative). We assume that the indicative inhibition of NR inference is the second main factor behind the limited NR-hood of Slavic languages and we believe that it is not coincidental that all of B&G's examples against NR in Slavic are based on indicative embedded clauses.

The indicative/subjunctive difference is orthogonal to our main interests and would probably merit a paper in its own right, so we will only shortly indicate what a possible approach could look like. One option is that strict NPIs are better licensed by the subjunctive than by the indicative mood, because the subjunctive is known to be more transparent for cross-clausal phenomena (see Progovac 1993 a.o.). Translating this into semantics, it was observed (see Villalta 2008) that only the subjunctive mood transfers alternatives from the embedded clause to the matrix predicate, whereas the indicative mood blocks them. Since the inference licensing of strict NPIs requires the projection of alternatives from the embedded clause, it would follow that the subjunctive mood is compatible with strict NPIs, while the indicative mood is not.

#### 4. CONCLUSION

In this article, we discussed an experiment targeting neg-raising in Czech. Contrary to B&G's (2009) claims, we argued that NR exists in Czech (a Slavic language without articles). We demonstrated that strict NPIs are more acceptable under NR predicates than under non-NR predicates, which follows naturally if Czech has a class of NR predicates. It is not clear to us how B&G (2009) could explain this contrast.

Somewhat surprisingly, we observed a contrast between clausemate negation and NR predicates with respect to strict NPI licensing. However, we explained these data while maintaining that NR-hood exists in Czech. In particular, we argued that the data follow under the scalar theory of NR (Romoli, 2012, 2013) and the anti-additive licensing condition for strict NPIs.

Several issues were discovered in our experiment and remain open. One is the observed effect of the mood on the licensing of strict NPIs. Another is the difference between 'even one' and 'until' and their interaction with licensing environments. Issues like these reveal that more than 50 years after Fillmore brought neg-raising to linguists' attention (Fillmore, 1963), neg-raising still does not have the last word.

#### References

- ABUSCH, Dorit (2002) "Lexical Alternatives as a Source of Pragmatic Presuppositions." *Semantics and Linguistic Theory (SALT)* 12, 1–19 <http://semanticsonline.org/pragmatics/protected/abusch.pdf>.
- BARTSCH, Renate (1973) "'Negative Transportation' gibt es nicht." *Linguistische Berichte* 27, 1–77.
- BOŠKOVIĆ, Željko/Jon GAJEWSKI (2009) "Semantic Correlates of the NP/DP Parameter." *Proceedings of NELS* 39, 121–134.

- CHIERCHIA, Gennaro (2004) “Scalar Implicatures, Polarity Phenomena, and the Syntax/Pragmatics Interface.” In: A. Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures*. Oxford: Oxford University Press, 39–103.
- CHIERCHIA, Gennaro (2013) *Logic in Grammar: Polarity, Free Choice, and Intervention*. Vol. 2. Oxford: Oxford University Press.
- COLLINS, Chris, and Paul M. POSTAL (2014) *Classical NEG Raising: An Essay on the Syntax of Negation*. Cambridge, MA: MIT Press.
- DOČEKAL, Mojmír (2014) “Intervention Effects: Negation, Universal Quantifiers, and Czech.” In: M. Ludmila/J. Veselovská (eds) *Complex Visibles Out There*. Olomouc: Palacký University, 369–387.
- DOČEKAL, Mojmír (2015) *Czech Negation from the Formal Perspective*. Munich: Lincom.
- FILLMORE, Charles J. (1963) “The Position of Embedding Transformations in a Grammar.” *Word* 1, 208–231.
- GAJEWSKI, Jon Robert (2005) *Neg-Raising: Polarity and Presupposition*. PhD Dissertation. Massachusetts Institute of Technology. <http://gajewski.uconn.edu/papers/thesis.pdf>.
- GAJEWSKI, Jon Robert (2007) “Neg-Raising and Polarity.” *Linguistics and Philosophy* 30/3, 289–328.
- GAJEWSKI, Jon Robert (2011) “Licensing Strong NPIs.” *Natural Language Semantics* 19, 109–148.
- HINTIKKA, Jaakko (1962) *Knowledge and Belief*. Ithaca, NY: Cornell University Press.
- HOMER, Vincent (2008) “Disruption of NPI Licensing: The Case of Presuppositions.” *Proceedings of SALT* 18, 429–46.
- HORN, Laurence R. (1989) *A Natural History of Negation*. Chicago: University Of Chicago Press.
- LADUSAW, William A. (1979) *Polarity Sensitivity as Inherent Scope Relations*. PhD Dissertation. University of Texas, Austin.
- LAKOFF, Robin (1969) “A Syntactic Argument for Negative Transportation.” *Proceedings of CLS* 5, 149–157.
- PROGOVAC, Ljiljana (1993) “Subjunctive: The (Mis)Behavior of Anaphora and Negative Polarity.” *The Linguistic Review* 10, 37–59.
- ROMOLI, Jacopo (2012) *Soft but Strong. Neg-Raising, Soft Triggers, and Exhaustification*. PhD Dissertation. Harvard University.
- ROMOLI, Jacopo (2013) “A Scalar Implicature-Based Approach to Neg-Raising.” *Linguistics and Philosophy* 36/4, 291–353.
- VILLALTA, Elisabeth (2008) “Mood and Gradability: An Investigation of the Subjunctive Mood in Spanish.” *Linguistics and Philosophy* 31/4, 467–522.
- ZWARTS, Frans (1998) “Three Types of Polarity.” In: F. Hamm/E. W. Hinrichs (eds), *Plurality and Quantification*. Dordrecht: Kluwer, 177–238.

Summary  
EXPERIMENTAL EVIDENCE FOR NEG-RAISING IN SLAVIC

Most research studying neg-raising focuses on English. Two notable exceptions are Bošković & Gajewski (2009) and Dočekal (2014) who discuss neg-raising in Slavic. In contrast to Bošković & Gajewski (2009), Dočekal (2014) argues that predicates of intention and obligation pass standard tests for neg-raisers in Czech. This article discusses new experimental data that provide additional evidence for the existence of neg-raising in Slavic languages, in particular, in Czech. The experiment that is conducted to test neg-raising predicates consists of an acceptability judgment task and an inference task. Sixty native speakers of Czech participated in the study. The results of the experiment are interpreted in Romoli's scalar theory of neg-raising (Romoli 2012, 2013). We claim that neg-raising exists in Czech, and argue that strict negative polarity items are more acceptable under neg-raising predicates than under non-neg-raising predicates.

**Keywords:** formal semantics, implicatures, neg-raising, Slavic, Czech

Povzetek  
EKSPERIMENTALNI DOKAZI ZA DVIGANJE NIKALNICE  
V SLOVANSKIH JEZIKIH

Večina raziskav s področja dviganja nikalnice se osredinja na angleščino. Prispevka Boškovića in Gajewskega (2009) ter Dočekala (2014) sta izjema na tem področju, saj obravnavata dviganje nikalnice v slovanskih jezikih. V nasprotju z izsledki Boškovića in Gajewskega (2009) Dočekal (2014) trdi, da predikati, ki izražajo namen in obvezo, dovoljujejo dviganje nikalnice v češčini. Pričujoči članek obravnava izsledke eksperimentalne raziskave, ki dodatno dokazujejo obstoj dviganja nikalnice v slovanskih jezikih, natančneje v češčini. Eksperiment, s katerim smo opazovali predikate z dviganjem nikalnic, sestavljata naloga s sodbami o sprejemljivosti in naloga iz sklepanja. Sodelovalo je šestdeset domačih govorcev češčine. Rezultati raziskave so obravnavani z vidika skalarne teorije o dviganju nikalnic avtorja Jacopa Romolija (Romoli 2012, 2013). V članku trdimo, da dviganje nikalnici v češčini obstaja, in zagovarjamo trditev, da so strogi k nikalnosti usmerjeni izrazi bolj spremenljivi v predikatih z dviganjem nikalnice kot v predikatih brez dviganja nikalnice.

**Ključne besede:** formalna semantika, implikatura, dviganje nikalnice, slovanski jeziki, češčina





## THE RELATIONSHIP BETWEEN (IN)ALIENABLE POSSESSION AND THE (THREE POTENTIAL) FORMS OF POSSESSED NOUNS IN HUNGARIAN\*\*\*

### 1. INTRODUCTION

In Hungarian, there are some nouns which have two different stems in the possessive paradigm. *Gyapja* ‘wool.Poss.3Sg’, for instance, is an inflected version of *gyapjú* ‘wool’. This has an inalienable interpretation since wool is an inalienable part of a sheep. However, if the wool is considered to belong to someone else, for instance, a shepherd, an alternative inflected form is used to express this alienable interpretation: *gyapjúja* ‘wool.Poss.3Sg’. In a similar fashion, the noun *ablak* ‘window’ also has an inalienable possessive form and an alienable one: *ablak-a* (that of a house) and *ablak-j-a* (say, that of a distributor). Section 2 scrutinizes this phenomenon, on the basis of which we classify possessed nouns into four groups (2.1), and base several generalizations on our observations concerning them (2.2).

It is shown in Section 3 that some productive Hungarian deverbal nominalizers, illustrated in (1) below, provide data relevant to the topic. Nominal constructions derived by means of these inevitably contain a possessed form of the noun head with shorter or longer variants of the possessedness suffix *-(j)A*, on the one hand, and on the other, a possessor with a thematic role designated in the derivational relationship (Laczkó 2000: 307-310; Alberti and Farkas to appear). In the case of  $T_{EV}$ -noun constructions, for instance, a shorter possessed form, claimed to indicate inalienability, is accompanied by a Theme possessor (1a), whilst in the case of  $T_{TH}$ -noun and *HATNÉK*-noun constructions, a longer possessed form is accompanied by an Agent-like possessor (1b-c).<sup>1</sup> Since the Agent is held to stand in a non-intrinsic, that is, alienable, relationship with the verb,

\* juttasusi@gmail.com

\*\* alberti.gabor@pte.hu

\*\*\* We are grateful to OTKA NK 100804 (*Comprehensive Grammar Resources: Hungarian*) for their financial support. The present scientific contribution is dedicated to the 650th anniversary of the foundation of the University of Pécs, Hungary.

1 As is also illustrated by the translations given in (1),  $T_{EV}$ -noun constructions denote complex events,  $T_{TH}$ -noun constructions refer to human Theme participants of complex events, while a *HATNÉK*-noun construction denotes a desire or urge concerning the realization of an event. For further discussion on *HATNÉK*-noun constructions, see Farkas and Alberti (to appear).



in contrast to the intrinsic relationship between verbs and their Themes (Kratzer 1996), the association of the longer possessed form with the less intrinsic semantic relationship is in harmony with den Dikken's (2015: 138) thesis based on a linguistic universal proposed by Haspelmath (2008), according to which longer possessive forms express alienable possession.

(1) Forms of possessives in some deverbal nominal constructions in Hungarian<sup>2</sup>

- |    |   |                           |                         |
|----|---|---------------------------|-------------------------|
| a) | <i>Vendel tegnap</i>  | <i>likvidál-t-á-val</i>   | [T <sub>EV</sub> -noun] |
|    | Vendel yesterday.Adj  | liquidate-T-Poss.3Sg-Ins  |                         |
|    | ‘with Vendel <sub>Theme</sub> (having been) liquidated yesterday’   |                           |                         |
| b) | <i>Vendel tegnap</i>  | <i>likvidál-t-ja</i>      | [T <sub>TH</sub> -noun] |
|    | Vendel yesterday.Adj  | liquidate-T-Poss.3Sg      |                         |
|    | ‘the person whom Vendel <sub>Agent</sub> liquidated yesterday’  |                           |                         |
| c) | <i>Vendel ebéd után</i>   | <i>való</i>               |                         |
|    | Vendel lunch after  | be.Part                   |                         |
|    | <i>beszélget-hetnék-je</i>  | <i>/ ásítóz-hatnék-ja</i> | [HATNÉK-noun]           |
|    | talk-HATNÉK-Poss.3Sg  | / yawn-HATNÉK-Poss.3Sg    |                         |
|    | ‘[Vendel’s <sub>Agent</sub> desire to talk] / [Vendel’s <sub>Partial_Agent</sub> urge to yawn] after lunch’ |                           |                         |

## 2. FORMS OF THE POSSESSEE

### 2.1 Four Groups of Nouns in Respect of Possessed Forms

The 3<sup>rd</sup> person singular possessedness suffix *-(j)A* has the following four allomorphs (in the case of singular possessives): *-ja*, *-je*, *-a*, and *-e*, distributed partly on the basis of vowel harmony, and partly on the basis of the following mysterious phenomenon, which attracts much attention in the literature. There are nouns which can appear more or less readily both with *-jA* and *-A* essentially depending on the alienable or inalienable semantic character of the possessive structure (e.g., Kiefer 1985, 2000: 201), as is illustrated by the minimal pair in (2a-a’) below. As the stem of the noun that bears the possessedness suffix may also appear in two different forms, illustrated by the often-quoted minimal pair in (5b-b’), nouns can be divided into four groups with respect to their potential alternative (3Sg) possessed forms. The relevant data is presented below in the series of examples in (2-5).

What is at stake is the verification in Hungarian of a straightforward generalization by Haspelmath (2008) according to which languages tend to express alienable possession by means of morphologically richer forms than inalienable possession. This can be done either by verifying that the component *-j-* itself has a morphemic status responsible for the expression of alienability inside the possessedness suffix *-(j)A* (den

2 In (1) and throughout the whole paper the following six-degree scale of grammaticality judgments, given in Broekhuis–Keizer–den Dikken (2012: viii), is used: \*: unacceptable, \*\*: relatively acceptable compared to \*, \*\*: intermediate or unclear status; ? : marked: not completely unacceptable or disfavored form; (°): slightly marked, but probably acceptable.

Dikken 2015: 131) or by interpreting morphological richness in some less trivial way (in (5b-b’), the expression of alienability is claimed to be transferred from *-j-* to *-á-* by den Dikken (2015: 141–142)).

Let us start the overview by considering the distribution of grammaticality judgments in the group of nouns with a single stem but with a phonotactically permitted alternation between the forms *-jA* and *-A* of the possessedness suffix *-(j)A* (2). As in all example pairs in (2-5), the possessive structures in the primed examples are evaluated as expressions of unquestionably alienable relationships while those in the primeless ones as those of inalienable relationships, or at least of types of relationship which can be regarded as encoded in Hungarian as inalienable on the basis of analogous examples.

Part-whole relations form the trivial basis of alienable possessive structures (2a,b,d,e) with body parts as a distinguished subset (2b,e).<sup>3</sup> Of these examples, in the inalienable constructions (2a,b,d), the *-jA* variants are fully unacceptable, while in the corresponding alienable possessive structures (2a’,b’,d’), the *-jA* variants are more or less marked but not unacceptable. In the latter case, the *-A* variants are also more or less marked (but still acceptable). As a similar distribution of grammaticality judgments can be observed in the minimal pair (2c-c’), rulers of nations can be considered to be encoded in language as inalienable parts of their nations.

(2) Inalienable/alienable forms of possessed nouns: I. Basic data, in which the difference between the variants can be regarded as *-j-* insertion

- |   |   |
|---|---|
| <p>a) <i>a ház ablak-(*)a</i><br/> the house window-Poss.3Sg<br/> ‘the window of the house’</p> | <p>a’) <i>a világ legjobb ablak-??(j)a<sup>4</sup></i><br/> the world best window-Poss.3Sg<br/> ‘the world’s best window’</p> |
| <p>b) <i>Ili talp-(*)a</i><br/> Ili sole-Poss.3Sg<br/> ‘Ili’s sole’</p>                         | <p>b’) <i>a világ legbűdösebb talp-?(?)a</i><br/> the world most_smelly sole-Poss.3Sg<br/> ‘the world’s most smelly sole’</p> |

3 Note that there is no difference between the four groups in the (semantic) respect that all contain body parts; see (2b,e), (3a-d), (4a-c), and (5c).

4 All of our alienable examples in (2-5) follow the pattern [possessor + superlative adjective + possessed noun] in order to guarantee that the alienable interpretation is achieved in a highly uniform manner. We are aware of the fact that there are also other constructions guaranteeing alienability as the [possessor + body part] construction in medical contexts and the [classifier as a possessed noun] construction (e.g., *?pohárja egye euró a sörnek* ‘glass.Poss.3Sg one euro the beer.Dat’ (‘a glass of beer costs one euro’). According to our first observations, there are slight (but fairly speaker-dependent) differences in grammaticality judgments between the different types of alienable construction. This suggests that (in)alienability is not a dichotomy but a scalar category. It goes beyond the scope of this paper to investigate this global aspect of the problem of forms of possessed nouns as well as to extend the investigation to forms of possessed nouns in plural and in non-third person. As for this latter problem, a noun like *ablak* ‘window’ (2a), for instance, has no alternative possessed forms in first person singular (*ablakom* ‘my window’ is the only form), in contrast to such nouns as *gyapjú* ‘wool’ (5c), which has a separate alienable form *gyapjóm* ‘my wool’ besides the inalienable form *gyapjam*.

- |   |  |
|---|--|
| c) <i>a németek császár-<sup>(?)j</sup>a</i><br>the German.Pl kaiser-Poss.3Sg<br>'the Germans' kaiser'  | c') <i>minden idők legifjabb császár-<sup>(?)j</sup>a</i><br>all time.Pl youngest kaiser-Poss.3Sg<br>'the youngest kaiser of all time'     |
| d) <i>az egyetem bölcsészkar-<sup>(*)j</sup>a</i><br>the university fac._of_hum.-Poss.3Sg<br>'the faculty of humanities<br>of the university' | d') <i>a világ legjobb bölcsészkar-<sup>(?)j</sup>a</i><br>the worldbest fac._of_hum.-Poss.3Sg<br>'the world's best faculty of humanities' |
| e) <i>Ili kar-<sup>(?)j</sup>a</i><br>Ili arm-Poss.3Sg<br>'Ili's arm'   | e') <i>a világ legerősebb kar-<sup>(?)j</sup>a</i><br>the worldstrongest arm-Poss.3Sg<br>'the world's strongest arm'                       |
| f) <i>az oroszok cár-<sup>(?)j</sup>a</i><br>the Russian.Pl tzar-Poss.3Sg<br>'the tzar of Russians'   | f') <i>minden idők legifjabb cár-<sup>(?)j</sup>a</i><br>all time.Pl youngest tzar-Poss.3Sg<br>'the youngest tzar of all time'             |

The last two examples, in which a body part (2e-e') and a sort of ruler (2f-f') are referred to, do not satisfy the above-sketched distribution of grammaticality judgments, since both kinds of interpretation can be expressed exclusively by the *-jA* variants. The homophonous forms (obviously belonging to the two different lexical items 'faculty' and 'arm') presented in (2d) and in (2e), thus, show different patterns of grammaticality judgments, in spite of the fact that both express part-whole relations.

In the series of examples in (3), a few phonotactic rules of Hungarian are illustrated which exclude the simultaneous occurrence of the *-jA* and *-A* variants in possessed nouns.

A noun ending in a vowel (3a-b'), for instance, has no *-A* variant while a noun ending in *-s* (pronounced as the consonant in the English word *ash*) (3c) has no *-jA* variant (see Rebrus 2014: 387-390; this exclusion also holds for all other sibilants as a very strong but somewhat speaker-dependent tendency). The consonant combination shown in (3d), however, disprefers *-A*. As for deciding the precise set of such consonants and consonant combinations, this morphophonological task (together with accompanying methodological questions) requires future research.

(3) Inalienable/alienable forms of possessed nouns: II. One (potential) form for phonotactic reasons

- |  |  |
|--|--|
| a) <i>Ili vesé-<sup>(?)j</sup>e</i><br>Ili kidney-Poss.3Sg<br>'Ili's kidney' | a') <i>a világ legnagyobb vesé-<sup>(?)j</sup>e</i><br>the world biggest kidney-Poss.3Sg<br>'the world's biggest kidney'           |
| b) <i>Ili boká-<sup>(?)j</sup>a</i><br>Ili ankle-Poss.3Sg<br>'Ili's ankle'   | b') <i>a világ legszebb boká-<sup>(?)j</sup>a</i><br>the world most_beautiful ankle-Poss.3Sg<br>'the world's most beautiful ankle' |
| c) <i>Ili has-<sup>(*)j</sup>a</i><br>Ili belly-Poss.3Sg<br>'Ili's belly'    | c') <i>a világ legnagyobb has-<sup>(*)j</sup>a</i><br>the world biggest belly-Poss.3Sg<br>'the world's biggest belly'              |

- |  |   |
|--|---|
| d) <i>Ili comb-*(<sup>✓</sup>j)a</i><br>Ili thigh-Poss.3Sg<br>'Ili's thigh'                      | d') <i>a világ legszebb comb-*(<sup>✓</sup>j)a</i><br>the world most_beautiful thigh-Poss.3Sg<br>'the world's most beautiful thigh' |
| e) <i>a csavar any*(<sup>(?)</sup>á)ja</i><br>the screw mother.Poss.3Sg<br>'the nut of the bolt' | e') <i>a szerelő legnagyobb any*(<sup>(?)</sup>á)ja</i><br>the mechanic biggest mother.Poss.3Sg<br>'the mechanic's biggest nut'     |

The minimal pair presented in (3e-e') above is of special interest to us, given that the noun *anya* 'nut (of a bolt)', which belongs to the vowel-ending subgroup shown in (3a-b) is a polysemic counterpart of *anya* 'mother' presented in (5b-b') below, just mentioned above as an example of nouns having two stems. This pair thus patterns with the pair of homophonous nouns presented in (2d-e') above in behaving differently in respect of accepting *-jA/-A* variants.

The third group consists of nouns with an alternative (idiosyncratic) possessed form from that which can be derived on-line from the nominative form via adding *-jA* or *-A* (4). Such nouns, thus, potentially have three 3Sg possessed forms. However, both the inalienable and alienable meanings are very much preferably expressed by the idiosyncratic variant and the *-A* variant is fully unacceptable. In the case of the noun *gyomor* 'stomach', for instance, the inalienable meaning can be expressed only by means of the idiosyncratic variant *gyomra*, providing a fully acceptable possessive structure (4a). The alienable meaning can be expressed by means of either the idiosyncratic form or the *-jA* variant, though both resulting structures are highly marked (4a'). The minimal pair in (4e-e') illustrates such an extreme preference for the idiosyncratic variant that this can readily express both kinds of meaning, with the two other potential forms providing fully unacceptable possessive structures.

(4) Inalienable/alienable forms of possessed nouns: III. An idiosyncratic form coexists with an "on-line created" form derived by means of *-jA* from the nominative version (while a form derived by means of *-A* is phonotactically possible but not acceptable)

- a) *Ili gyomra / \*gyomorja / \*gyomora*  
Ili stomach.Poss.3Sg  
'Ili's stomach'
- a') *a világ legnagyobb ??gyomra / ??gyomorja / \*gyomora*  
the world biggest stomach.Poss.3Sg  
'the world's biggest stomach'
- b) *a sas karma / \*karomja / \*karoma*  
the eagle claw.Poss.3Sg  
'the claw of the eagle'
- b') *a világ legélesebb ??karma / \*\*karomja / \*karoma*  
the world sharpest claw.Poss.3Sg  
'the world's sharpest claw'

- c) *Ili körme / \*körömje / \*köröme*  
 Ili nail.Poss.3Sg  
 ‘Ili’s nail’
- c’) *a világ legélesebb<sup>??</sup> körme / \*körömje / \*köröme*  
 the world sharpest nail.Poss.3Sg  
 ‘the world’s sharpest nail’
- d) *Pécs főtere / \*főtérje / \*főtére*  
 Pécs main\_square.Poss.3Sg  
 ‘the main square of Pécs’
- d’) *a világ legnagyobb<sup>(?)</sup> főtere / \*főtérje / \*főtére*  
 the world biggest main\_square.Poss.3Sg  
 ‘the world’s biggest main square’
- e) *Pécs egyik tere / \*térje / \*tére*  
 Pécs one\_of square.Poss.3Sg  
 ‘a square of Pécs’
- e’) *a világ legnagyobb tere / \*térje / \*tére*  
 the world biggest square.Poss.3Sg  
 ‘the world’s biggest square’

In the fourth group, the nouns have an idiosyncratic possessed form but the ending of the nominative form excludes either the *-A* variant (5a-c’,e-e’) or the *-jA* variant (5d-d’) for the same phonotactic reasons as was discussed in connection with the examples presented in (3) above. It can be observed that the inalienable meanings can be associated only with the idiosyncratic variants (see the primeless examples in (5a-d)). The alienable meanings, on the other hand, are only associated with the variants based on the nominative form (see the corresponding primed examples). The minimal pair presented in (5e-e’) with the noun *falva* ‘village’ is somewhat exceptional with respect to the inalienable meaning. Presumably this is due to the quite archaic character of the idiosyncratic variant *falva*: in present-day Hungarian, the nominative-form-based variant *faluja* is almost as acceptable as the idiosyncratic variant (NB: it is even questionable whether the possessive structure presented in (5e) is encoded as an inalienable relationship in language). As for the alienable meaning tested in (5e’), it is unequivocally the nominative-form-based variant that expresses the alienable meaning (even more preferably than in the case of the acceptability pattern typical of the corresponding variants in (5a’,b’,c’,d’)).

(5) Inalienable/alienable forms of possessed nouns: IV. An idiosyncratic form coexists with an on-line created form whilst phonotactics prohibit *-A/-jA* alternation

- |  |   |
|--|---|
| a) <i>a ház teteje / *tetője</i><br>the house roof.Poss.3Sg<br>‘the roof of the house’ | a’) <i>a cég legjobb teteje / <sup>(?)</sup>tetője</i><br>the firm best roof.Poss.3Sg<br>‘the firm’s best roof’ |
|--|---|

- |  |   |
|--|---|
| b) <i>Ili anyja</i> / * <i>anyája</i><br>Ili mother.Poss.3Sg<br>'Ili's mother'   | b') <i>a világ legjobb</i> * <i>anyja</i> / <sup>(?)</sup> <i>anyája</i><br>the world best mother.Poss.3Sg<br>'the world's best mother'                   |
| c) <i>a juh gyapja</i> / * <i>gyapjúja</i><br>the sheep wool.Poss.3Sg<br>'the wool of the sheep'   | c') <i>a cég legjobb</i> * <i>gyapja</i> / <sup>(?)</sup> <i>gyapjúja</i><br>the firm best wool.Poss.3Sg<br>'the firm's best wool'                        |
| d) <i>a tűz parazsa</i> / * <i>parázsa</i><br>the fire glow.Poss.3Sg<br>'the glow of fire'   | d') <i>a világ legforróbb</i> <sup>??</sup> <i>parazsa</i> / <sup>?</sup> <i>parázsa</i><br>the world hottest glow.Poss.3Sg<br>'the world's hottest glow' |
| e) <i>a zsellérek</i> <sup>(?)</sup> <i>falva</i> / <sup>?</sup> <i>faluja</i><br>the cottar.Pl village.Poss.3Sg<br>'the village of cottars' | e') <i>a világ legjobb</i> * <i>falva</i> / <sup>?</sup> <i>faluja</i><br>the world best village.Poss.3Sg<br>'the world's best village'                   |

## 2.2 Generalizations

In what follows, the observations about the four groups of nouns are summarized (Table 1) and generalized using the terminology defined in (6) below, which makes it possible to formulate the generalizations in a simple and elegant form (7).

### (6) Definition of three kinds of possessed variants

- a) *Possessed variant 1* (v1):  
on-line created as [nominative form of the noun + *-jA*], unless the relevant phonotactic rules of Hungarian prohibit this;  
otherwise, [nominative form of the noun + *-A*].<sup>5</sup>
- b) *Possessed variant 2* (v2):  
on-line created as [nominative form of the noun + *-A*] if the relevant phonotactic rules of Hungarian permit both this variant and the [nominative form of the noun + *-jA*] variant  
(NB: v2 is defined in a way that it is inevitably different from v1).
- c) *Possessed variant 3* (v3):  
acceptable (idiosyncratic) historical form of the noun, different from those referred to in v1 and v2, if extant.

It must be noted that certain speakers refuse variant 1 forms in most cases, saying that they sound very artificial (e.g., *ablakja* 'its window'; cf. the generally accepted v2 variant *ablaka*). This phenomenon may be regarded as a kind of hypercorrection: the speakers in question are convinced that the given variants violate certain rules they learned, in spite of the fact that they have never been taught such rules. Certain variant 3 forms are also problematic for some speakers because they consider them unacceptably

5 The variant *hasa* 'his/her/its belly' (3c), for instance, counts as a (potential) variant v1 according to (6a), since the form *hasja* is (considered to be) phonotactically excluded even from the set of competing *potential* variants.

archaic (e.g., *disznaja* ‘his/her pig’; cf. the “modern” v1 variant *disznója*). It also generally holds for all examples in (2-5) except for those marked as fully acceptable or fully unacceptable that they show quite high speaker-dependent variation.

In Table 1, the four quarters correspond to the four series of examples in (2-5).

The (simplest) bottom right quarter presents the grammaticality judgments given in (3). In this group, v1 has no potential alternative, since there is no idiosyncratic variant (v3) and phonotactic rules exclude another nominative-form-based variant (v2). Furthermore, what is worth noting is that, in all cases in this group, v1 can always readily express the inalienable meaning as well as the alienable one.

The top right quarter presents the grammaticality judgments given in (5). In the corresponding group of nouns, only v3 and v1 are “in competition”, indicated in the corresponding heading as ‘{3, 1}’, since v2 is excluded for phonotactic reasons. As was observed in (5), in this group, v3 can readily express the inalienable meaning (indicated by the formula ‘{3, 1} → 3’ in the table) and v1 can be almost as readily associated with the alienable meaning (‘{3, 1} → 1’), but not *vice versa*. Thus the available potential variants differentiate the two kinds of meanings in a plausible way. The primary inalienable meaning belongs to the idiosyncratic variant v3 whilst the alienable meaning, calculated in the given context on the basis of some kind of mental conceptual network (Alberti and Farkas to appear, 2.1.1.2.2),<sup>6</sup> is expressed by the variant which can be calculated automatically, that is, by v1, referred to as the primary on-line created variant in (6a). Note that the simplified formula in (7a) below ([v1~a]) refers to this latter relationship between the alienable meaning and the on-line created form, which can be regarded as a generalization over den Dikken’s (2015: 131, 141-142) *-j*-insertion in certain cases (2a,c,d), and *-á*-insertion (5b) in other cases, plus some further morphological differences for which den Dikken (2015) does not account (5a,c,d,e).<sup>7</sup>

6 In such a mental network, it must be calculated that, say, the possessive structure *my house* can refer not only to default relationships such as my owning the given house, and/or my living there, but also arbitrarily expanded relationships such as my being the homeless person who inspects the garbage cans of the house or my being the agent whose task is to make the residents fill in some questionnaire.

7 It is clear that this group can be characterized not by a difference between the competing possessive variants manifesting itself in a certain sound-size morpheme (or sequence of sounds) but by the phonetically highly varied and unpredictable difference between an automatically producible potential variant v1 and an idiosyncratic variant v3 existing for historical reasons. The variant v1 *tető-je* ‘its roof’, for instance, is different from the v3 variant *tete-je* in the quality and the length of the stem-final vowel (see also the minimal pairs *ajtó-ja/ajta-ja* ‘its door’, *disznó-ja/diszna-ja* ‘its pig’, *tüdő-je/tüde-je* ‘its lung’). Relative to the v1 variant *gyapjú-ja* ‘its wool’, however, the v3 variant *gyapj-a* does not contain a stem-final vowel of another quality but it lacks the stem-final vowel and (hence) this form gets the *-j*-less version of the suffix *-(j)A* (somewhat similar examples with other stem-final vowels: *anyú-ja/any-ja* ‘its mother’, *apú-ja/ap-ja* ‘its father’). The minimal pair *parázs-a/parazs-a* ‘its ember’ exemplifies the case when there is a difference (chiefly) in the length of a stem-internal vowel (also see *darázs-a/darazs-a* ‘its wasp’), while the pair *falv-ja/falv-a* ‘its village’ illustrates the type of difference based on the phenomenon often referred to as *v*-insertion in synchronic descriptions (also see *tetű-je/tetv-e* ‘its louse’).



Table 1: Acceptability of different variants of possessed forms depending on phonotactic and historical factors

-A / -jA IDIOSYN	YES				NO			
	INALIENABLE		ALIENABLE		INALIENABLE		ALIENABLE	
YES	{3, 2, 1} → 3		{3, 2, 1} → 3		{3, 1} → 3		{3, 1} → 1	
	3	1	3	2	3	1	3	1
	✓gyomra	*gyomorja	??gyomra	??gyomorja	✓teteje	*tetője	*teteje	(?)tetője
	✓karma	*karomja	??karma	*karomja	✓anyja	*anyája	*anyja	(?)anyája
	✓körme	*kőrömje	??körme	*kőrömje	✓gyapja	*gyapjúja	*gyapja	(?)gyapjúja
	✓főtere	*főtérje	(?)főtere	*főtérje	✓parazsa	*parázsa	??parazsa	?parázsa
	✓tere	*térje	✓tere	*térje	(?)falva	?faluja	*falva	✓faluja
	NO	{2, 1} → 2		{2, 1} → 1		{1} → 1		{1} → 1
2		1	2	1	1		1	
✓ablaka		*ablakja	??ablaka	?ablakja	✓veséje		✓veséje	
{2, 1} → 2		{2, 1} → 2						
2		1	2	1	✓bokája		✓bokája	
✓talpa		*talpja	?talpa	??talpja	✓hasa		✓hasa	
✓b.kara		*b.karja	(?)b.kara	?b.karja	✓combja		✓combja	
✓császára		*császársa	(?)császára	?császársa	(?)anyája <sub>nut</sub>		(?)anyája <sub>nut</sub>	
{2, 1} → 2		{2, 1} → 2						
2		1	2	1				
*kara		✓karja	*kara	?karja				
*cára		✓cárja	*cára	✓cárja				

The top left quarter of Table 1 presents the grammaticality judgments given in (4). In the corresponding group of nouns, it could be theoretically possible that all the three variants be in competition, but, as can be observed in (4), v2 cannot express either the inalienable meaning or the alienable one. The systematic unacceptability of v2 is indicated in the corresponding heading by crossing out this variant (see the notation ‘{3, 2, 1}’ in the top left quarter of the table). The table does not present the uniformly fully unacceptable data. A generalization can be formulated which holds for all types of data that v2 and v3 mutually exclude each other; see (7e) below (\*[v3 & v2]). Hence there is no noun with three more or less acceptable possessed forms (7e’) (\*[v1 & v2 & v3]). Another straightforward consequence of the mutual exclusion between v2 and v3 is that if a noun has two possessed forms, one of them is v1 (7e’’).



Variant v2 excluded, hence, both in the top right quarter and in the top left quarter: the same two variants “remain in competition”. However, the outcome in the two cases is different: while in the top right quarter, the two variants differentiate the two kinds of meanings, as is formulated in (7a-a’) ([v1~a, v2~i, v3~i]), in the top left quarter, both the inalienable meaning and the alienable meaning are significantly more readily expressed by v3 (cf. (7c’): [v3~a → \*v1, \*v2]). What is formulated in (7a-a’) is a (plausible) strategy that functions only in certain domains of nouns (see the two specially framed domains with dark rims in Table 1) and not a universal generalization valid for all Hungarian nouns.

The bottom left quarter of the table, in which (in the absence of idiosyncratic (v3) alternatives) the two nominative-form-based variants v1 and v2 are in competition, shows the most eclectic picture. This picture is a reflection of the great variety according to which certain data pattern with those in the top right quarter in associating different forms with the two kinds of meanings (7a-a’) whilst other data pattern with those in the top left quarter in associating the same forms with the two kinds of meanings, and moreover, both v1 and v2 can serve as this dominant form (in the case of different nouns, of course; cf. (7c-c’): [v1~i → \*v2, \*v3], [v2~a → \*v1, \*v3]). Thus in this quarter, both v1 and v2 are associated with either the inalienable or the alienable meaning (in the case of different nouns), as is registered in (7b) below ([v1~i / v2~a / v3~a]); nevertheless, it never occurs that, in the case of one and the same noun, the alienable meaning is expressed by v2 while the inalienable one by v1. This restriction, which holds for all data in all the four quarters, is formulated in (7d) as follows: if different variants are associated with the two kinds of meaning (see the two specially framed domains with dark rims in Table 1), the “sequence number” (1, 2, and 3 given in (6) above) of the variant belonging to the inalienable meaning must be greater than that of the variant belonging to the alienable meaning; the opposite association is excluded. Note that this generalization can completely cover all the data with no exception due to its formulation in which cases of equation are also accepted (included in the relation ‘k≥n’ in (7d)): such cases cover the nouns with a single acceptable possessed form (see the other four domains framed with light rims in Table 1).

(7) Generalizations on the (somewhat hidden) relationship between v1 *versus* v2,v3 and alienability (a) *versus* inalienability (i)

a) v1~a	on-line created: morphophonologically ~ semantically
a’) v2~i, v3~i:	(a-a’): partial tendency as a good point of departure
b) v1~i / v2~a / v3~a	there are such counterexamples
c) [v1~i → *v2, *v3]	v1 as dominant variant
c’) [v2~a → *v1, *v3]	v2 as dominant variant
c’’) [v3~a → *v1, *v2]	v3 as dominant variant
d) [v <sub>k</sub> /v <sub>n</sub> ~ i/a → k≥n]	inalienable/alienable for each domain
e) *[v3 & v2]	variants excluding each other
e’) *[v1 & v2 & v3]	all the three variants cannot appear simultaneously
e’’) [v <sub>k</sub> & v <sub>n</sub> (k>n) → n=1]	of two variants, one is v1

All in all, although the data in (2-5) suggest a close relationship between the automatically calculable (on-line created) variant v1 and a context-dependent alienable meaning (7a), and, parallel to this, between the other two variants and inalienable meaning (7a'), either kind of meaning can be expressed by any variant (7b).

The hypothesized asymmetry of semantic affiliation between v1 *versus* v2 and v3, beyond the fact that the natural strategy formulated in (7a-a') explicitly prevails in certain domains of nouns, also prevails in the other domains "vacuously" and "implicitly" in the following sense. By 'vacuously' we mean that there is no domain in which v1 expresses the inalienable meaning with v2 or v3 expressing the alienable meaning (7d). As for 'implicit' manifestations of the asymmetry in question, the (c)-constraints formulate them by claiming that if a variant can express the opposite kind of meaning relative to its basic character given in (7a-a'), then, in the case of the same noun, it will express (at least as readily) the other kind of meaning (the one that *ab ovo* suits it), too. In such cases, the given variant is referred to as a *dominant* one.

The constraints in (7e-e'') formulate restrictions on the coexistence or, on the contrary, dominance of the three variants which do not follow from the foregoing.

Let us consider a few detailed instances or consequences of the (e)-restrictions (already discussed above). If, for instance, the idiosyncratic variant v3 expresses the inalienable meaning and phonotactics does not exclude the construction of a variant v2, the latter will be fully unacceptable as an expression of either the same inalienable meaning or the alienable meaning. The latter part of this claim is in harmony with the spirit of (7a-a') while the former part can be regarded as a unicity condition: it is needless to express the same kind of meaning in two or more ways. Therefore, practically if a noun has two (more or less acceptable) possessed forms, then the alienable meaning is expressed by v1 (7e'') and the inalienable meaning either by v2 or by v3, exclusively; it follows that there is no noun with three different more or less acceptable possessed forms (7e').

### **3. RELATIONSHIP BETWEEN FORMS OF POSSESSED DEVERBAL NOMINALS AND THE THEMATIC CHARACTER OF THEIR POSSESSORS**

Possessive structures of (complex-eventuality-related) derived nouns with thematic possessors fit well in the system functioning according to the constraints presented in (7) above – through placing the given types of derived noun in the appropriate quarters of Table 1; see Table 2 below. As is illustrated in (8) below (which demonstrates the entire system, only a part of which was shown in (1) in the Introduction), in the case of complex-eventuality-related derived nouns, the possessor always corresponds to a designated input argument, that is, it is always a thematic argument (see Alberti and Farkas to appear). Thematic arguments appearing as possessors are of distinguished relevance because the Agent is held to stand in a non-intrinsic relationship with the verb (Kratzer 1996), which can plausibly be considered to be related to alienability, in contrast to the intrinsic (hence, inalienable) relationship between verbs and their Themes.

(8) Productive deverbal nominalizers in Hungarian

- a) *Vendel tegnapi likvidál-t-á-val* [T<sub>Ev</sub>-noun]  
 Vendel yesterday.Adj liquidate-T-Poss.3Sg-Ins  
 ‘with Vendel<sub>Theme</sub> (having been) liquidated yesterday’
- b) *Vendel tegnapi*  
 Vendel yesterday.Adj  
*el-rohan-ás-a / likvidál-ás-a* [ÁS-noun]  
 away-run-ÁS-Poss.3Sg / liquidate-ÁS-Poss.3Sg  
 ‘the fact that yesterday [Vendel<sub>Agent</sub> ran away] / [Vendel<sub>Theme</sub> was liquidated]’
- c) *Vendel tegnapi likvidál-ó-ja* [Ó-noun]  
 Vendel yesterday.Adj liquidate-ó-Poss.3Sg  
 ‘the person who liquidated Vendel<sub>Theme</sub> yesterday’
- d) *Vendel tegnapi likvidál-t-ja* [T<sub>Th</sub>-noun]  
 Vendel yesterday.Adj liquidate-T-Poss.3Sg  
 ‘the person whom Vendel<sub>Agent</sub> liquidated yesterday’
- e) *Vendel ebéd után való*  
 Vendel lunch after be.Part  
*beszélget-hetnék-je / ásítóz-hatnék-ja* [HATNÉK-noun]  
 talk-HATNÉK-Poss.3Sg / yawn-HATNÉK-Poss.3Sg  
 ‘[Vendel’s<sub>Agent</sub> desire to talk] / [Vendel’s<sub>Partial Agent</sub> urge to yawn] after lunch’
- e’) *Vendel i-hatnék-\*(j)a / e-hetnék-?(j)e*  
 Vendel drink-HATNÉK-Poss.3Sg / eat-HATNÉK-Poss.3Sg  
 ‘Vendel’s<sub>Agent</sub> desire to drink / eat’
- e’’) *Vendel ásítóz-hatnék-\*(j)a / tüszent-hetnék-?(j)e*  
 Vendel yawn-HATNÉK-Poss.3Sg / sneeze-HATNÉK-Poss.3Sg  
 ‘Vendel’s<sub>Partial Agent</sub> urge to yawn / sneeze’

Since complex-eventuality-related derived nouns are inherently on-line created, they have no idiosyncratic possessed forms, so they cannot appear in the top two quarters of Table 1, but must be sorted in the bottom quarters according to phonotactic factors. ÁS-nouns (8b) and ó-nouns (8c) must obviously be placed in the bottom right quarter, since ÁS-nouns end in *-s* (cf. *has\*(j)a* ‘its belly’ in (3c)) and ó-nouns end in a vowel (cf. *vesé-\*(j)e* ‘its kidney’ and *boká-\*(j)a* ‘its ankle’ in (3a-b)), so for them to have v2 is excluded by the phonotactic rules of Hungarian (NB: ÁS-nouns have only *-A* variants and ó-nouns have only *-jA* variants).<sup>8</sup>

8 As is illustrated by the translations given in (8b-c), ÁS-noun constructions denote complex events, while ó-noun constructions primarily refer to Agent participants of complex events (on Instrument/Location-denoting ó-noun constructions, see subsection 1.3.1.3 in Alberti and Farkas (to appear)). As was mentioned in footnote 2, T<sub>Ev</sub>-noun constructions denote complex events, T<sub>Th</sub>-noun constructions refer to human Theme participants of complex events, and HATNÉK-noun constructions denote a desire or urge concerning the realization of an event.

Complex-event-based  $\tau$ -nouns (see (8a,d)), however, follow the (7a-a') strategy in the bottom left quarter of the classification of possessed-noun forms in Table 1. That is,  $\tau_{Ev}$ -nouns (8a), which tend to have “inalienable” Theme possessors (and never prototypical Agents under any circumstances), have v2 possessed forms, while  $\tau_{Th}$ -nouns (8d) have v1 possessed forms, since their possessors are not Themes (though  $\tau_{Th}$ -noun constructions as a whole *refer to* human Themes). Since possessors of HATNÉK-nouns (8e) in the corresponding verbal argument structures are not Themes, either, but prototypical Agents or Agent-like participants who have partial control over bodily/sound emission, possessed forms of HATNÉK-nouns are – correctly – predicted to be variants v1. As the comparison between the grammaticality judgments associated with the minimal pairs in (8e'-e'') above shows, there *are* differences between the potential possessed HATNÉK-noun variants. However, these differences are not due to the completely or partially agentive character of the possessor (compare (8e') with (8e'')) but to such phonetic factors as the HATNÉK-noun form's demand for velar  $-(j)a$  or palatal suffixes  $-(j)e$ . HATNÉK-nouns requiring palatal suffixes, in contrast to those requiring velar suffixes, accept  $-j$ -less possessed forms to a certain extent (such forms have somewhat marginal grammaticality), also intensively depending on dialectal differences.

Table 2: The classification of  $-A/-jA$  forms of possessed complex-event(uality)-related derived nouns in Table 1 (depending on the thematic character of possessors)

3? \ 2?	YES	NO															
YES	{3, 2, 1} → 3/3	{3, 1} → 3/1															
NO	{2, 1} → 2/2	<table style="margin-left: auto; margin-right: auto; border-collapse: collapse;"> <tr> <td colspan="3" style="text-align: center;">{2, 1}</td> </tr> <tr> <td colspan="3" style="text-align: center;">↓</td> </tr> <tr> <td style="text-align: center;">2</td> <td style="text-align: center;">/</td> <td style="text-align: center;">1</td> </tr> <tr> <td style="text-align: center;"><math>\tau_{Ev}</math></td> <td></td> <td style="text-align: center;"><math>\tau_{Th}</math></td> </tr> <tr> <td colspan="3" style="text-align: center;">HATNÉK</td> </tr> </table>	{2, 1}			↓			2	/	1	$\tau_{Ev}$		$\tau_{Th}$	HATNÉK		
	{2, 1}																
↓																	
2	/	1															
$\tau_{Ev}$		$\tau_{Th}$															
HATNÉK																	
	{2, 1} → 1/1	<table style="margin-left: auto; margin-right: auto; border-collapse: collapse;"> <tr> <td colspan="3" style="text-align: center;">{1}</td> </tr> <tr> <td colspan="3" style="text-align: center;">↓</td> </tr> <tr> <td style="text-align: center;">1</td> <td style="text-align: center;">/</td> <td style="text-align: center;">1</td> </tr> <tr> <td style="text-align: center;">ÁS</td> <td></td> <td style="text-align: center;">ÁS</td> </tr> <tr> <td style="text-align: center;">Ó</td> <td></td> <td style="text-align: center;">.</td> </tr> </table>	{1}			↓			1	/	1	ÁS		ÁS	Ó		.
{1}																	
↓																	
1	/	1															
ÁS		ÁS															
Ó		.															

#### 4. CONCLUSION

We argue that den Dikken's (2015) hypothesis concerning the existence of a morpheme  $-j-$  in Hungarian responsible for the expression of alienability must be generalized into (and should be replaced with) a system of more abstract and conditional claims (given in (7)) in order to account for all the relevant data (Section 2), including deverbal

nominals with possessors carrying different thematic roles (Section 3). In this global picture, den Dikken’s (2015) hypothesis appears as a (plausible) strategy that functions in a single domain of nouns (with competing variants v1 with *-jA* and *-j*-less v2 variants; see the lower specially framed domain with dark rims in Table 1), while in other domains, the *ab ovo* association of the *-jA* variants with alienability manifests itself in more hidden forms that (i) the alienable *-jA* variant (or rather, what is defined as variant v1 in (6a)) is opposed to a (phonetically varied) group of alternative idiosyncratic (v3) possessee variants (see footnote 8), or (ii) there is a dominant possessee variant (7c-c’), which simply suppresses the other potential variants (blocking or covering the differentiation according to (in)alienability).

## References

- ALBERTI, Gábor/Judit FARKAS (to appear) “Derivation of Nouns.” In: G. Alberti/T. Laczkó (eds), *Nouns and Noun Phrases*. Amsterdam: AUP.
- BROEKHUIS, Hans/Evelien KEIZER/Marcel DEN DIKKEN (2012) *Syntax of Dutch – Nouns and Noun Phrases, Vol I-II*. Amsterdam: AUP.
- DEN DIKKEN, Marcel (2015) “The morphosyntax of (in)alienably possessed noun phrases: The Hungarian contribution.” In: K. É. Kiss, B. Surányi and É. Dékány (eds), *Approaches to Hungarian 14*. Amsterdam: John Benjamins, 121–145.
- FARKAS Judit/Gábor ALBERTI (to appear) “The Hungarian HATNÉK-noun expression: a hybrid construction.” In: A. Bondaruk/A. Bloch-Rozmej (eds), *Constraints on structure and derivation in syntax, phonology and morphology*. Frankfurt am Main: Peter Lang.
- HASPELMATH, Martin (2008) *Syntactic universals and usage frequency, 3: Alienable vs. inalienable possessive constructions*. Handout, Leipzig Spring School on Linguistic Diversity.
- KIEFER, Ferenc (1985) “Natural morphology.” *Acta Linguistica Hungarica* 35, 85–105.
- KIEFER, Ferenc (2000) *Jelentésmélet*. Budapest: Corvina.
- KRATZER, Angelika (1996) “Severing the External Argument from its Verb.” In: J. Rooryck/L. Zaring (eds), *Phrase Structure and the Lexicon. Studies in Natural Language and Linguistic Theory*, Volume 33. Dordrecht: Kluwer, 109–137.
- LACZKÓ, Tibor (2000) “Az ige argumentumszerkezetét megőrző főnévképzés.” In: Kiefer F. (ed.), *Strukturális magyar nyelvtan 3. Morfológia*. Budapest: Akadémiai Kiadó, 293–407.
- REBRUS, Péter (2014) “Miért nincs j?” In: A. Benő, E. Fazakas and E. Kádár (eds), *... hogy legyen a víznek lefolyása*. Kolozsvár: Erdélyi Múzeum-Egyesület Kiadó, 383–401.

Summary  
THE RELATIONSHIP BETWEEN (IN)ALIENABLE POSSESSION  
AND THE (THREE POTENTIAL) FORMS OF POSSESSED NOUNS  
IN HUNGARIAN

The paper gives a thorough insight into the system of possible forms of (in)alienably possessed nouns in Hungarian. Its point of departure is the group of [Nominative + -j- +A] possessive forms the stem of which has an alternative (morphologically “shorter”) possessive form; such longer possessive forms are claimed to express alienable possession (see den Dikken 2015). We point out that Hungarian deverbal nominals—and especially the groups of T-nouns—play an interesting role in this system via the thematic character of their possessors (given the obvious connection between alienable possession and external argumenthood, on the one hand, and inalienable possession and internal argumenthood, on the other).

**Keywords:** (in)alienable possession, Hungarian, possessedness suffix (j)A, deverbal nominals, thematic roles

Povzetek  
RAZMERJE MED (NE)ODTUJLJIVO SVOJILNOSTJO IN  
(TREMI POTENCIALNIMI) OBLIKAMI POSEDOVANIH  
SAMOSTALNIKOV V MADŽARŠČINI

V članku podrobno predstavimo sistem možnih oblik (ne)odtujljivo posedovanih samostalnikov v madžarščini. Izhajamo iz svojilnih oblik tipa [Imenovalnik + -j- +A], katerih osnova pozna tudi alternativno (morfološko “krajšo”) svojilno obliko; tovrstne daljše svojilne oblike naj bi izražale odtujljivo svojilnost (glej den Dikken 2015). Med drugim izpostavimo, da imajo izglagolski samostalniki v madžarščini – še posebej skupina T-samostalnikov – v omenjenem sistemu zanimivo vlogo, ki izhaja iz tematske narave njihovih posedovalcev (ob upoštevanju očitne povezave med odtujljivo svojilnostjo in vlogo zunanjega argumenta na eni strani ter neodtujljivo svojilnostjo in vlogo notranjega argumenta na drugi strani).

**Ključne besede:** (ne)odtujljiva svojilnost, madžarščina, pripona posedovanega (j)A, izglagolski samostalniki, udeleženske vloge





## YES OR NO, OR HOW TO ANSWER A NEGATIVE QUESTION

### 1. INTRODUCTION

The aim of this paper is to investigate the status of particles in answers to negative questions. A *yes-no question* is a question that asks to identify whether P or not P is true (Hamblin 1973), and is minimally answered by an answering particle. Answers to *yes-no* questions have given rise to a large amount of literature since Pope (1972), who gives a detailed description of the types of answers across languages. She distinguishes two major systems: truth-based systems, in which the particles confirm or disconfirm the true value presupposed by the question, like in Japanese in (1), and polarity-based systems, in which particles express positive or negative polarity, like in Spanish in (2).

- (1) a) *Kimi tsukareteiru?*      *Hai. // Iie.*      (Jap)<sup>1</sup>  
 you tired                      yes    no  
 ‘Are you tired?’              ‘Yes (I am).’ // ‘No (I am not).’
- b) *Kimi tsukareteinai?*      *Hai (tsukareteinai). // Iie (tsukareteiru desu).*  
 you tired-neg                yes (tired-neg)    no (tired be)  
 ‘Aren’t you tired?’        ‘I am not’. // ‘I am.’
- (2) a) *¿Enviaste una carta a Paul?*      *Sì. // No.*      (Sp)  
 ‘Did you send a letter to Paul?’      ‘Yes.’ // ‘No.’
- b) *¿No enviaste una carta a Paul?*      *No. // Sì.*  
 ‘Didn’t you send a letter to Paul?’      ‘No.’ // ‘Yes, I did.’

The same particles are used to answer both positive and negative questions, although some languages “reinforce” the positive particle to negative questions, e.g. the conjunction *but* before *yes* in Czech, see (3), and some languages use a specific particle for positive answers to negative questions, e.g. *si* instead of *oui* in French, see (4).

- (3) a) *Poslal jsi Pavlovi dopis?*      *Ano. // Ne.*      (= 2a) (Cz)  
 sent be.2sg to-Paul letter      yes    no  
 ‘Did you send Paul a letter?’
- b) *Neposlal jsi Pavlovi dopis?*      *Ne. // Ale ano.*      (= 2b) (Cz)  
 neg-sent be.2sg to-Paul letter      no      but yes  
 ‘Didn’t you send Paul a letter?’

\* hana.gruet-skrabalova@univ-bpclermont.fr

1 This example is from my Japanese informant.



- (4) a) *Est-il venu?*                      *Oui. // Non.*                      (Fr)  
       is-he come                              yes    no  
       ‘Has he come?’
- b) *N’ est-il pas venu?*              *Non. // \*Oui. / Si.*  
       neg-did-he not come              no     yes    yes  
       ‘Hasn’t he come?’

Although the behaviour of particles has been described for many languages (see Laka 1990; Martins 1994; Holmberg 2001; Holmberg 2012; Farkas 2010; Krifka 2012 a.o.), it has been hardly observed that particles could be used in both ways within a same language. So, Holmberg (2001; 2013) notes that negative questions with *not* in English can be confirmed by both *yes* and *no*, see (5), because *not* is ambiguous between sentential and VP-negation. In other words, he argues that *yes* only confirms negative questions when *not* is interpreted as VP-negation, thus when the question is rather affirmative than truly negative.

- (5) a) *Is John not coming? No, he is not.* (sentential Neg) / *Yes, he is not.* (VP-Neg)  
       b) *Isn’t John coming? No, he isn’t.* (sentential Neg) / *\*Yes, he isn’t.* (\*VP-Neg)

This paper shows that negative questions in Czech can be both confirmed and disconfirmed by *yes* or *no*, see (6), which makes the particles potentially ambiguous. The negation is expressed by the negative prefix on the finite verb, cf. (3) above.

- (6) *Rodiče nejsou doma?*              *Ne. / Ano. // Ne. / Ale ano.*  
       parents neg-are home              no    yes    no    but    yes  
       ‘Aren’t the parents at home?’    ‘They are not.’ // ‘They are.’

The use of the particles is however not free. It is argued that it depends on the interpretation of the sentential negation in the question, which can be either true or expletive (cf. Brown and Franks 1995 for Russian negative questions). This semantic distinction is furthermore tightly linked to the syntactic position of the negation, according to which we can distinguish between negative interrogative clauses and negative declarative clauses used as questions (cf. Gunlogson 2001). An analysis in terms of absolute and relative polarity (Farkas 2010) is then proposed to account for the mixed behaviour of answering particles: particles express absolute polarity in answers to interrogative questions, whose polarity is open (Holmberg 2001), and relative polarity in answers to declarative questions, whose polarity has been already specified.

The paper is structured as follows. Section 2 describes two types of answers to negative questions in Czech, showing a mixed behaviour of answering particles. Section 3 deals with negative questions; several pieces of evidence are discussed in order to show that we have to distinguish between questions with expletive negation and questions with true negation and that this distinction accounts for the distribution of answering particles. Section 4 focuses on expression of the polarity in the answers; it is

proposed that particles express absolute or relative polarity depending on the polarity of the question. Section 5 sums up the paper.

## 2. TWO TYPES OF ANSWERS TO NEGATIVE QUESTIONS

To show that both answering particles in Czech can confirm and disconfirm a negative question, we need to specify that Czech (like many other languages, e.g. Basque, Finish, Portuguese, Irish) can reply a question by using the finite verb of the question, either alone or in combination with a particle, see (7). The positive verb always indicates a positive answer and the negative verb a negative answer. Verbal answers will therefore be used throughout the paper to clearly state the polarity of the answer.

- (7) a) *Poslal jsi Pavlovi dopis?* (Ano) poslal. // (Ne) neposal.  
 sent be.2sg to-Paul letter yes sent no neg.sent  
 ‘Did you send Paul a letter?’
- b) *Neposal jsi Pavlovi dopis?* (Ne) neposal. // (Ale ano) poslal.  
 neg-sent be.2sg to-Paul letter no neg.sent but yes sent  
 ‘Didn’t you send Paul a letter?’

Let us look again at the negative question in (6), repeated in (8) and completed by verbal answers (in brackets) that can accompany the particles:

- (8) *Rodiče nejsou doma?* A: *Ne (nejdou).* // *Ale ano (jsou).*  
 parents neg-are home no neg-are but yes are  
 B: *Ano (nejdou).* // *Ne (jsou).*  
 yes neg-are no are  
 ‘Aren’t the parents at home?’ ‘They are not.’ // ‘They are.’

The puzzle we are dealing here is that the question ‘Aren’t the parents at home?’ can be answered in two ways, which I call type A and type B answers and which are summarized in table 1 below. In type A answers, *ne* means ‘they are not’, while *ano* means ‘they are’. In contrast, in type B answers, *ne* means ‘they are’, while *ano* means ‘they are not’. The particles appear, thus, to be ambiguous; they may also combine with both negative and positive verb. The interpretation of each answer is however clearly given by the verb, meaning that the combination of the particle and the verb cannot be considered as redundant.<sup>2</sup>

---

2 Speakers actually often prefer verbal answers to answering particles alone.

Table 1: Two types of answers to negative questions

	Negative answers	Positive answers
Type A answers	<b>ne + neg-V</b>	<b>(ale) ano + verb</b>
Type B answers	<b>ano + neg-V</b>	<b>ne + V</b>

Example (8) seems to indicate that A and B answers are always equally available, but this is not true. First, in oral production, the particles in B answers are better emphasized and separated from the verb, which confirms that they bear another information than the verb (i.e. they do not indicate the absolute polarity of the answer, see section 4):

- (9) *Jan nemluví francouzsky? ?Ne, mluví. // NE – mluví.*  
 John neg-speaks French no he-speaks NO he-speaks  
 ‘John doesn’t speak French?’

Furthermore, B answers are generally judged unacceptable as replies to questions containing an initial verb, meaning that the question’s form matters for the choice of the answer (see section 3 for more details):

- (10) *Neposlal jsi mu ten dopis? \*Ano (neposlal). / Neposlal. // \*Ne (poslal). / Poslal.*  
 neg-sent be.2sg to-him the letter yes (neg-sent) neg-sent no (sent) sent  
 ‘Didn’t you send him the letter?’

Importantly, however, B answers can be used as replies to negative questions with negative polarity items like *ani* (not-one/not-any), see (11). Since *any* must be licensed by sentential negation (see section 3.3 for more on NPIs), negative questions confirmed by B answers must contain sentential negation. They, therefore, differ from negative questions with *not* in English that allow confirmation by *yes* only with VP-negation reading, see (5) above.

- (11) *Jan \*(ne)pozval ani jednoho spolužáka? Ano (ani jednoho). // Ne (nějakého pozval).*  
 John (neg-)invited no one schoolmate yes no one no (he) some invited  
 ‘Didn’t John invite any of his schoolmates?’

The difference between A and B answers cannot thus be due to the contrast between sentential and VP-negation, but it must nevertheless be linked to the negation. The hypothesis I explore in the following section is that in some negative questions, negation loses its negative force and behaves like expletive negation. Negative questions with expletive negation behave like positive questions, whose polarity I assume to be open (following Holmberg 2001). Thus, they are answered by type A answers, in which particles indicate absolute (positive or negative) polarity. In contrast, negative questions

with true negation behave like negative assertions, whose polarity is fixed. Thus, they are answered by type B answers, in which particles indicate relative polarity, i.e. (dis)agreement with the polarity of the question.

### 3. NEGATIVE QUESTIONS

The main claim of this section is that in some questions, negation is not true (from the truth conditional point of view), but expletive. The term ‘expletive’ (or ‘pleonastic’) negation usually refers to a phonologically overt negative morpheme that lacks negative semantic content (see Espinal 1992 for Romance; Brown & Franks 1995 for Slavic). It typically appears after lexical items like *before* or *until*, and in constructions with verbs like *to fear* or *to doubt*, see (12). Contrary to true negation, expletive negation cannot license NI-phrases in Slavic, compare (13a) and (13b).<sup>3</sup> For Brown & Franks (1995: 262), “canonical pleonastic negation (in Slavic) consists of a NegP with either an empty or vacuous specifier position. The head position is filled with *ne* [...], but there is no Negation Operator, the bearer of the semantics, to give the sentence negative force”.

- (12) a) *Il n'est pas arrivé.* (ne...pas: true Neg)<sup>4</sup> (Fr)  
 he neg-is not arrived  
 ‘He hasn’t arrived.’
- b) *Il faut finir avant qu’il n’arrive.* (ne: expletive Neg)  
 it must finish before that-he neg-arrives  
 ‘We have to finish before he arrives.’
- c) *Je suis sûr que nous ne sommes pas en retard.* (ne: true Neg)  
 I am sure that we neg are not late  
 ‘I’m sure that we are not late.’
- d) *Je crains que nous ne soyons en retard.* (ne: expletive Neg)  
 I fear that we neg are late  
 ‘I am afraid that we are late.’
- (13) a) *Nikdo nepřišel.* (ne-: true Neg) (Cz)  
 nobody neg-come  
 ‘Nobody came.’
- b) *Bojím se, aby někdo / \*nikdo nepřišel pozdě.* (ne-: expletive Neg)  
 I-fear refl. that somebody / nobody neg-come late  
 ‘I am afraid that somebody might come late.’

More interestingly, Brown & Franks (1995) and Abels (2002) observe that negative questions with the interrogative particle *li* in Russian do not license NI-phrases,

3 See section 3.4 for licensing of NPIs in Czech.

4 In (12a) and (12c), *ne* can be omitted.

which would suggest that they also contain expletive negation. However, the same questions license Genitive of Negation. Moreover, questions without particle *li* license both NI-phrases and Genitive of Negation. To account for this asymmetry, Brown & Franks (1995: 266) claim that “certain independently motivated grammatical principles, such as Rizzi’s (1990) Relativized Minimality, conspire to prevent negation from having negative force [...] in Russian Yes/No questions”. In their analysis, negation in *li*-questions is forced to be pleonastic, because a Yes/No operator in Spec-CP prevents *ne* to be licensed by the Negation Operator (adjoined to CP) when *ne* (cliticized on V) moves to C. I pursue here the view that negation can be rendered expletive by the interrogation. In other words, the interrogative operator blocks the negative operator to bind *ne* when Verb(Neg)-moves to C, so that *ne* cannot retain its negative force. I will show that despite the absence of Genitive of Negation in Czech, the phenomenon of expletive negation is relevant and allows to account for the distribution of A and B answers to negative questions.

### 3.1 Word Order

Main questions in Czech do not contain an overt interrogative C. Still, assuming that syntactic clause-type (declarative, interrogative, exclamative) and illocutionary act (assertion, questioning, exclamation) can be dissociated (Beysade & Marandin 2006), we can distinguish two types of questions. Czech is a SVO language in which interrogative clauses are formed by V-fronting (and *wh*-fronting in case of *wh*-questions), see (14). Interrogative clauses are typically used as questions. But declarative clauses with (not only) SVO order can also be used to express questioning. In the following discussion, I will show that negative interrogatives and negative declarative clauses used as questions differ with respect to negative presupposition, combinability with evidential and modal adverbials, licensing of NPIs, and also the answers they receive.

- (14) a) *Marie dala dětem koláčky. / \*Dala Marie dětem koláčky.* (declarative)  
 Mary.nom given children.dat cakes.acc /given Mary.nom children.dat cakes.acc  
 ‘Mary gave children some cakes.’
- b) *Dala Marie dětem koláčky? / Komu dala Marie koláčky?* (interrogative)  
 given Mary.nom children.dat cakes.acc /who.dat given Mary.nom cakes.acc  
 ‘Did Mary give children some cakes?’ / ‘To whom did Mary give some cakes?’

### 3.2 Negative Presupposition

Negative interrogatives (with non-focal intonation<sup>5</sup>) are used to elicit information, to solicit an opinion, or to make a polite request for action (see Gunlogson 2001 for detailed distribution). They do not convey any presupposition. Questions in (15) are understood as open questions that the speaker uses to find out whether the doctor is or is not in his office, and whether the addressee did or did not take a wrong road. They can only be answered by type A answers.

5 See ex. (21) for focal intonation.

- (15) a) *Není doktor v ordinaci? Ne, není. / \*Ano, není. // Ano, je. / \*Ne, je.*  
 neg-is doctor in office no neg-is yes is yes is no is  
 ‘Isn’t the doctor in his office?’ ‘He is not.’ // ‘He is.’
- b) Q: *Kam jdete? Nespletli jste si náhodou cestu?*  
 where go.2pl neg-mistaken be.2pl refl. accidentally road  
 ‘Where are you going? Didn’t you accidentally take the wrong road?’
- A: *Ne, nespletli. / \*Ano, nespletli. // Ano, spletli. / \*Ne, spletli.*  
 no neg-mistaken yes neg-mistaken yes mistaken no mistaken  
 ‘We didn’t.’ // ‘We did.’

On the contrary, negative declarative questions are typically used when the speaker expects a negative answer or when (s)he wishes to express an emotional response (surprise, disagreement, etc.). They convey a negative presupposition like negative declarative clauses. Questions in (16) are typically used when the speaker expects the doctor not to be in his office or if he is surprised that the addressee did not find the right way. Type B answers are then used to confirm or disconfirm this negative bias.

- (16)
- a) *Doktor není v ordinaci? Ano, není. // Ne, je.*  
 doctor neg-is in office yes neg-is no is  
 ‘The doctor isn’t in his office?’ ‘He isn’t.’ // ‘He is.’
- b) *Jak je to možné? Vy jste nenašli správnou cestu? Ano, nenašli. // Ne, našli.*  
 how is it possible you be.2pl neg-found right way yes neg-found no found  
 ‘How is it possible? You didn’t find the right way?’ ‘We didn’t.’ // ‘We did.’

However, the biased interpretation in (16) is not obligatory, which explains that these questions can be also answered by type A answers, as shown in (16’) (see section 3.4, ex. (24) for the analysis). The ambiguity of declarative questions can be avoided by using final rising intonation for neutral interpretation (like in interrogative clauses) and rising intonation followed by final fall for biased interpretation.

- (16’) a) *Doktor není v ordinaci? Ne, není. // Ale ano, je.*  
 doctor neg-is in office no neg-is but yes is  
 ‘The doctor isn’t in his office?’ ‘He isn’t.’ // ‘He is.’

### 3.3 Evidential and Modal Adverbs

The fact that only negative declarative questions can express a negative bias is confirmed by their compatibility with adverbs expressing modality or evidentiality, like *určitě* (‘of course’), *zřejmě* (‘apparently’), *jistě* (‘surely’), see (17). These adverbs are considered as bias markers and cannot appear in interrogative clauses in English either, see (18) (cf. Huddleston 1994).

- (17) a) *Pavel s tím určitě nesouhlasil? / #Nesouhlasil s tím Pavel určitě?*  
 Paul with that certainly neg-agreed / neg-agreed with that Paul certainly  
 ‘Paul did certainly not agree with that?’
- b) *Prosefoři ještě zřejmě neodešli? / #Neodešli ještě profesoři zřejmě?*  
 professors yet apparently neg-left / neg-left yet professors apparently  
 ‘Apparently, the professors haven’t yet left?’
- c) *Vy nejste pravděpodobně zdejší? / #Nejste pravděpodobně zdejší?*  
 you neg-are probably from-here / you-neg-are probably from-here  
 ‘You are probably not from here?’
- (18) a) *#Did they certainly agree with that?*  
 b) *#Have they apparently left?*  
 c) *#Are you probably from here?*

### 3.4 Negative Polarity Items (NPIs)

Three types of negative polarity items can be distinguished in Czech: NI-phrases (*nikdo* ‘nobody’, *nic* ‘nothing’, *žádný N*, ‘no N’, etc.), weak NPIs (*vůbec* ‘at all’, *sebemenší* ‘the slightest’) and strict NPIs (*ani jeden* ‘no one’), see Dočekal (2016). NI-phrases are only licensed by clausemate negation, i.e. they must be in the scope of sentential negation, see (19a). Weak NPIs are licensed in the context of some downward entailing operator (*yes-no* questions, conditionals, constituent negation etc., see Gajewski 2011). The strict NPI *ani jeden* is grammatical in contexts with local or superordinate negation, see (19b) and (19c), and in contexts with the preposition *bez* ‘without’.<sup>6</sup>

- (19) a) *Nikdo nepotkal / \*potkal v lese medvěda.*  
 nobody neg-met / met in forest bear  
 ‘Nobody met a bear in the forest.’
- b) *Ani jeden z nich se v lese neztratil / ztratil.*  
 no one of them refl. in forest neg-lost / lost  
 ‘Not one of them got lost in the forest.’
- c) *Nechce, aby se ani jeden ztratil v lese.*  
 he-neg-wants that refl. no one lost in forest  
 ‘He doesn’t want anyone to get lost in the forest.’

If we look at NI-phrases in questions, we can observe that negative declarative questions license NI-phrases as subject exactly like negative declarative clauses, while negative interrogatives do not, see (20). Moreover, negative interrogatives in (20a) can only be answered by type A answers, while B answers are possible in (20b). The

<sup>6</sup> It is also sensitive to truth conditions and to the pragmatic part of meaning, which explains that it is not acceptable in all downward entailing contexts (Dočekal 2016).

two contrasts are explained if interrogative clauses contain expletive negation and declarative clauses true negation, cf. (13) above.<sup>7</sup>

- (20) a) *Nepotkal \*nikdo / někdo v lese medvěda? Ne / \*Ano, nepotkal.*  
 neg-met nobody / somebody in forest bear no yes neg-met (he did not)  
 ‘Did anybody meet a bear in the forest?’ *Ale ano / \*Ne, potkal.*  
 but yes no met (he did)
- b) *Nikdo / někdo nepotkal v lese medvěda? Ano (nikdo) / Ne (někdo ho potkal).*  
 nobody / somebody neg-met in forest bear yes (nobody) no (somebody did)  
 ‘Nobody met/Somebody didn’t meet a bear in the forest?’

The contrast observed above needs to be specified in more detail. First, NI-phrases are ungrammatical in interrogative clauses in (20a) and (21a) below, but they are acceptable in (21b), where they are used with focal interpretation. Brown & Franks (1995) argue however that negation cannot be expletive with focal interrogation because the verb doesn’t raise to C (recall that expletive negation in questions is triggered in contexts with Verb(Neg)-raising to C). I claim that the verb in (21b) actually doesn’t rise to C but stays in T, whose specifier is occupied by a null expletive subject (because of the EPP feature). The focused subject stays in Spec-vP. The fact that an overt expletive subject (*von* ‘he’) can co-occur with the focused subject in (21c) makes this claim plausible.

- (21) a) *Nepotřebuje \*nikdo / někdo pomoci?*  
 neg-needs nobody / somebody help  
 ‘Does anyone need help?’
- b) *Nepotřebuje NIKDO pomoci?*  
 neg-needs nobody help  
 ‘Nobody needs help?’
- c) *Von nepotřebuje NIKDO pomoci?*  
 he-expl neg-needs nobody help  
 ‘Nobody needs help?’

Second, contrary to NI-phrases, strict NPIs are grammatical in both types of questions, see (22). This contrast parallels Russian *li*-questions in which negation does not license NI-phrases, but triggers Genitive of Negation (see above). I suggest, therefore, that strict NPIs are licensed before negation moves to the interrogative C, where it loses

7 The expletive status of negation can be supported by the behaviour of PPI. The PPI in the sentence is interpreted (as expected) only with the wide scope with respect to negation, while the PPI in the sentence (ii) with V(Neg)-raising is not. I thank Mojmir Dočekal for this observation.

(i) *Někdo nepotkal včera Karla?* = ‘is there a specific x who did not meet Karel?’  
 somebody neg-met yesterday Karel<sub>acc</sub>

(ii) *Nepotkal někdo včera Karla?* = ‘is there or isn’t there an x who met Karel’  
 neg-met somebody yesterday Karel<sub>acc</sub>



its negative force. In contrast, NI-phrases must stay in local relation with Negation throughout the derivation. Consequently, expletive negation cannot license NI-phrases at LF, see (23b).

- (22) a) *Nepotřebuje ani jeden z vás pomoci?*  
 neg-needs no one from you help  
 ‘Doesn’t anyone from you need help?’

- b) before Verb(Neg)-raising:  $\checkmark [_{TP} [_{NegP} Neg [_{vP} \dots ani \dots]]]$   
 c) after Verb(Neg)-raising:  $\checkmark Neg+C [_{TP} \dots [_{NegP} [_{vP} \dots ani \dots]]]$

- (23) a) before Verb(Neg)-raising:  $\checkmark [_{TP} [_{NegP} Neg [_{vP} \dots ni\text{-phrase} \dots]]]$   
 b) after Verb(Neg)-raising:  $*Neg+C [_{TP} \dots [_{NegP} [_{vP} \dots ni\text{-phrase} \dots]]]$

Finally, recall that negation in declarative questions can also have expletive interpretation. This suggests that the verb movement to C can be covert; consequently, we obtain type A answers to negative declarative questions:

- (24) *Doktor není v ordinaci?* → LF: *Není doktor v ordinaci? Ne (není).*  
 doctor neg-is in office? neg-is doctor in office no (neg-is)  
 ‘The doctor isn’t in his office?’ ‘He isn’t.’

### 3.5 Summary

Negative questions in Czech can be expressed by using either interrogative or declaratives clauses. These two types of clauses behave differently with respect to several properties (presupposition, adverbs, NPIs, see table 2) that can be explained if the sentential negation in interrogative clauses loses its negative force (thus becomes expletive) by virtue of its movement to the C. Importantly, the distribution of type A and B answers described in section 2 also follows from the distinction between true and expletive negation.

Table 2: Interrogative vs. declarative negative questions

	Negative interrogatives	Negative declarative questions	
Word order	VSO	SVO	
NI-phrases	*	✓	
Negative presupposition	No	Yes	No
Negation	Expletive	True	Expletive
Answers	A	B	A

#### 4. BACK TO ANSWERS

It is generally assumed (since Laka 1990) that answering particles are generated in the polarity projection ( $\Sigma P$ , PolP) in the CP domain, because they express positive or negative polarity. The concept of the polarity must be, nevertheless, made clearer in order to account for the mixed behaviour of answering particles in Czech.

Following Farkas (2010), I propose to distinguish between absolute and relative polarity or rather between absolute and relative value of the polarity. In declarative clauses, polarity is a feature that has two absolute values: positive [+] and negative [-]. The relation between the two polarity values can be called relative polarity. In the case of question-answer pairs, relative polarity indicates the relation between the polarity of the question [Q] and the polarity of the answer [A]. There are four possible relations: either [Q] and [A] have the same positive or negative value, or they have different values, one being positive and the other negative. The basic idea is that the particle *ano* ('yes') expresses the positive value [+], or (by default) the relation between two identical values: [+,+], and [-,-] = [+]. The particle *ne* ('no') expresses the negative value [-], or (by default) the relation between two different values: [+,-] and [-,+], = [-].

##### 4.1 Expressing Polarity in Answers to Interrogative Clauses

Positive interrogatives have open polarity (*x*, cf. Holmberg 2001), because they ask whether [P] or [not P], [P] corresponding to the positive polarity value and [not P] to the negative polarity value. The particles are minimal answers in that they only indicate the polarity value of the P, the P itself being presupposed (and elided). The polarity head of PolP receives its value by specifier-head agreement with the particle in its specifier:

- (25) *Chtěl<sub>x</sub> byste šálek čaje?* a) [<sub>PolP</sub> *Ano*<sub>[+]</sub> [<sub>Pol'</sub> [<sub>XP</sub>  $\emptyset$ ]]] (= *chtěl*)  
 wanted be.cond.2pl cup of-tea yes wanted  
 'Would you like a cup of tea?' 'I would like a cup of tea.'
- b) [<sub>PolP</sub> *Ne*<sub>[-]</sub> [<sub>Pol'</sub> [<sub>XP</sub>  $\emptyset$ ]]] (= *nechtěl*)  
 no neg-wanted  
 'I wouldn't like a cup of tea.'

Assuming that negation becomes expletive by virtue of Verb(Neg)-raising, negative interrogatives also have open polarity and therefore behave like positive interrogatives. Consequently, they are answered by type A answers:

- (26) *Nechtěl<sub>x</sub> byste šálek čaje?* a) [<sub>PolP</sub> *Ano*<sub>[+]</sub> [<sub>Pol'</sub> [<sub>XP</sub>  $\emptyset$ ]]] (= *chtěl*)  
 neg-wanted be.cond.2pl cup of-tea yes wanted  
 'Wouldn't you like a cup of tea?' 'I would like a cup of tea.'
- b) [<sub>PolP</sub> *Ne*<sub>[-]</sub> [<sub>Pol'</sub> [<sub>XP</sub>  $\emptyset$ ]]] (= *nechtěl*)  
 no neg-wanted  
 'I wouldn't like a cup of tea.'

To say that negative interrogatives behave as positive interrogatives with respect to negation does not mean that they are completely equivalent. They cannot, for instance, be used to initiate a line of inquiry or to raise an issue as open or unsettled, as shown in (27). On the contrary, they are willingly used as polite requests and to solicit advice or an opinion, see (28).<sup>8</sup>

- (27) a) *Máte děti? Pokud ano, chodí / #nechodí do školy?*  
 have.2pl children if yes they-go / neg-go to school  
 ‘Have you got children? If yes, do / #don’t they go to school?’
- b) *Co myslíte, bude / #nebude François Hollande znovu zvolen prezidentem?*  
 what think.2pl will / neg-will François Hollande again elected president  
 ‘What do you think, will / #will not François Holland be re-elected president?’  
*Může být, ale také nemusí.*  
 he-can be but also neg-must  
 ‘It could go either way.’
- (28) a) (v tramvaji) *Chcete / Nechcete pustit sednout?*  
 (in a tram) want.2pl / neg-want.2pl let sit  
 ‘Would(n’t) you like to take my seat?’
- b) *Co myslíš, mám / nemám se nechat ostříhat?*  
 what think.2sg I-have / neg-have refl. let hair-cut  
 ‘What do you think, should(n’t) I get my hair cut?’

As noted by Brown & Franks (1995), negative interrogatives must, thus, be endowed with some specific communicative and pragmatic value. Negative interrogatives in Czech are actually considered as more polite or less direct than the positive ones.<sup>9</sup> I suggest the following explanation for such a politeness effect.<sup>10</sup> A polar question asks the addressee to choose between two alternatives, positive or negative (P or not P), but refusing or saying *no* can be perceived as rude, and, therefore, difficult to express. By using the negative form of the question, the speaker presents the negative alternative as acceptable and consequently allows the addressee to express his refusal more easily. It is precisely because negation has lost its truth-conditional role that it can play such a role at the discourse level (see Groenendijk & Stokhof 1997).

8 A detailed semantic and pragmatic approach of these questions can be found in Reese (2006) and Krifka (2012).

9 See Leech (2014: 167) for negation as “a strategy to express a degree of polite indirectness”.

10 A similar effect can be observed with negative-raising. The sentences *Nemyslím, že s tím bude souhlasit* (‘I don’t think he will agree with that’) and *Myslím, že s tím nebude souhlasit* (‘I think he will not agree with that’) are semantically equal (i.e. the negation is interpreted in the embedded clause), but only the raised negation adds a politeness effect to the sentence interpretation.

## 4.2 Polarity in Answers to Declaratives Clauses

I have said above that declarative clauses have their polarity feature specified. The value of the polarity feature in negative declarative clauses is negative. Likewise, negative declarative clauses used as questions have a negative polarity value. That is the reason why they convey a negative presupposition (see section 3.1). The answer confirms or disconfirms this negative presupposition. Consequently, the particles only indicate relative polarity, i.e. relation between the polarity of the question and that of the answer.<sup>11</sup>

- (29) *Oni ten návrh nepřijali<sub>[-]</sub>?*      a) [<sub>PolP</sub> *Ano*<sub>[-,-]</sub> [<sub>Pol'</sub> [-] [<sub>XP</sub>  $\emptyset$ ]]] (= *nepřijali*)  
 they the proposal neg-accepted      yes      neg-accepted  
 ‘They didn’t accept the proposal?’      ‘They didn’t.’
- b) [<sub>PolP</sub> *Ne*<sub>[+,-]</sub> [<sub>Pol'</sub> [-] [<sub>XP</sub>  $\emptyset$ ]]] (= *přijali*)  
 no      accepted  
 ‘They did.’

Importantly, this analysis can be extended to positive questions conveying a positive presupposition, as in (30). But there will be no difference at surface between answers expressing absolute polarity and those expressing relative polarity. In both, (25) above and (30) below, the positive answer will be expressed by “*ano* + positive verb”, and the negative answer by “*ne* + negative verb”. The mixed behaviour of the particles can therefore be only observed with negative questions.

- (30) *Oni ten návrh přijali<sub>[+]</sub>?*      a) [<sub>PolP</sub> *Ano*<sub>[+,-]</sub> [<sub>Pol'</sub> [+] [<sub>XP</sub>  $\emptyset$ ]]] (= *přijali*)  
 they the proposal accepted      yes      accepted  
 ‘They accepted the proposal?’      ‘They did.’
- b) [<sub>PolP</sub> *Ne*<sub>[+,-]</sub> [<sub>Pol'</sub> [+] [<sub>XP</sub>  $\emptyset$ ]]] (= *nepřijali*)  
 no      neg-accepted  
 ‘They didn’t.’

Particles are, therefore, not ambiguous, but their felicitous use depends on the felicitous interpretation of the question. This can be resolved by using specific prosody schemes to indicate biased or focused interpretation. Finally, the proposed analysis predicts that the mixed behaviour of particles can be observed in languages in which negation raising to a particular C leads to its expletive interpretation and in which declarative clauses can be used as questions. This prediction seems to come true e.g. for Russian and Spanish (see Gruet-Skrabalova 2014), as shown in (31) and (32) respectively.<sup>12</sup> This issue is, however, out of the scope of this paper and must be left to future research.

11 *Ano* in (29) cannot be itself (without *ale* ‘but’) interpreted as “they did accept”. *Ne* is theoretically ambiguous, but in practice, the ambiguity will be resolved by the intonation of both question and answer (see section 3.2., ex. (16) and (16’), and section 3.4 ex. (24)).

12 These examples come from my Russian and Spanish informants.

- (31) a) *Ne priexali li roditeli domoj?*      *Net, ne priexali. // Da, priexali.*  
 neg came    part. parents at-home    no neg came    yes came  
 ‘Haven’t the parent come at home?’
- b) *Roditeli ne priexali domoj?*      *DA ne priexali. // NET, priexali.*  
 parents neg came    home    yes neg came    no came  
 ‘The parents haven’t come at home?’
- (32) a) *¿No están en casa los padres?*      *No, no están. // Si, están.*  
 neg are    at home the parents    no neg are    yes are  
 ‘Aren’t the parents at home?’
- b) *¿Los padres no están en casa?*      *¿SI, no están. // NO, están.*  
 the parents neg are    at home    YES neg are    //NO are  
 ‘The parents aren’t at home?’

## 5. CONCLUSION

In this paper, I have dealt with answers to negative *yes-no* questions, focusing on data from Czech. I have shown that answering particles can express both positive and negative answers to negative questions, but that their distribution depends on the semantic and syntactic properties of questions. I have argued that negation in questions loses its negative force when it moves to the interrogative head C and behaves thus as expletive negation. Consequently, I have distinguished between two types of negative questions: (i) negative interrogative clauses that contain an initial negative verb and expletive negation, and that do not convey a negative presupposition, and (ii) negative declarative clauses used as questions, that contain a non-initial negative verb and true negation, and that convey a negative presupposition. Following this syntactic and semantic distinction, I have shown that negative interrogatives receive the same answers as positive interrogatives (*yes* in case of positive answer and *no* in case of negative answer), while negative declaratives used as questions are confirmed by *yes* (corresponding to a negative answer) and denied by *no* (corresponding to a positive answer). I have explained this distribution by proposing that the particles express absolute polarity in answers to interrogative questions, whose polarity is open and must be fixed by the particle, and relative polarity in answers to declarative questions, whose polarity has been already specified.

## References

- ABELS, Klaus (2002) “Expletive (!) Negation.” In: J. Toman (ed.), *Proceedings of FASL 10*, 1–20.
- BEYSSADE, Claire/Jean-Marie MARANDIN (2006). “The Speech Act 97 Assignment Problem Revisited: Disentangling Speaker’s Commitment from Speaker’s Call on Addressee”. In: O. Bomani/ P. Cabredo-Hofherr (eds), *Empirical Issues in Syntax and Semantics 6*, <http://www.cssp.cnrs.fr/eiss6/>, 37–68.
- BROWN, Sue (1999) *The Syntax of Negation in Russian*. Stanford: CSLI Publications.

- BROWN, Sue/Steven FRANKS (1995) "Asymmetries in the scope of Russian negation." *Journal of Slavic Linguistics* 3/2, 239–287.
- DOČEKAL, Mojmír (2016) "Constituent negation, subjunctive and alternatives." Talk presented at The Olomouc Linguistics Colloquium (*Olinco*, June 2016), Palacký University Olomouc.
- ESPINAL, Maria T. (1992) "Expletive Negation and Logical Absorption." *Linguistic Review* 9, 333–358.
- FARKAS, Donka (2010) "The grammar of polarity particles in Romanian." In: A. Di Sciullo/ V. Hill (eds), *Edges, Heads, and Projection: Interface properties*, Amsterdam: John Benjamins, 81–124.
- FARKAS, Donka F./Kim B. BRUCE (2010) "On reacting to assertions and polar questions." *Journal of Semantics* 27, 81–118.
- GAJEWSKI, Jon (2011) "Licensing strong NPIs." *Natural Language Semantics* 19, 109–148.
- GROENENDIJK, Jeroen/Martin STOKHOF (1997) "Questions." In: J. van Benthem/ A. ter Meulen (eds), *Handbook of logic and language*, Amsterdam: Elsevier, 1055–124.
- GRUET-SKRABALOVA, Hana (2015) "Verbs and particles in answers to polar questions." In: G. Zybatow *et al.* (eds), *Slavic Grammar from a Formal Perspective*, Frankfurt am Main: Peter Lang, 197–215.
- GRUET-SKRABALOVA, Hana (2014) "Two kinds of minimal answers to yes-no question in Czech and Spanish." XI Congreso Internacional de Lingüística General, 21–23 May 2014, Universidad de Navarra, Pamplona. To appear in A. Gordejuela Senosiain *et al.* (eds), *Lenguas, lenguaje y lingüística. Contribuciones des de la Lingüística General*. Pamplona: Servicio de Publicaciones de la Universidad de Navarra.
- GUNLOGSON, Christine (2001) *True to form: Rising and Falling Declaratives as Questions in English*. PhD diss., Univ. of California at Santa Cruz.
- HAMBLIN, Charles L. (1973) "Questions in Montague English." *Foundations of Language* 10, 41–53.
- HOLMBERG, Anders (2001) "The syntax of *yes* and *no* in Finnish." *Studia linguistica* 55, 141–174.
- HOLMBERG, Anders (2012) "On the syntax of *yes* and *no* in English." *Newcastle Working Papers in Linguistics* 18, 52–72.
- HOLMBERG, Anders (2013) "The syntax of negative questions and their answers." In: N. Gato *et al.* (eds), *Proceedings of GLOW in Asia* 9, 1–18.
- HUDDLESTON, Rodney (1994) "The contrast between interrogatives and questions." *Journal of Linguistics* 30/2, 411–439.
- KRIFKA, Manfred (2013) "Response particles as propositional anaphors." *Proceedings of SALT* 23, 1–18.
- LAKA, Itziar (1990) *On the Syntax of Negation: On the nature of functional categories and projections*. PhD diss. Cambridge, MIT.
- LEECH, Geoffrey (2014) *The Pragmatics of Politeness*. New York: Oxford University Press.

- MARTINS, Ana (1994) "Enclisis, VP-deletion, and the nature of Sigma." *PROBUS* 6/2-3, 173–206.
- POPE, Emily N. (1972) *Questions and answers in English*. PhD. Diss. MIT.
- REESE, Brian J. (2006) "The Meaning and Use of Negative Polar Interrogatives." In: O. Bonami/ P. Cabredo Hofherr (eds), *Empirical Issues in Syntax and Semantics* 6, 331–354.

### Summary

#### YES OR NO, OR HOW TO ANSWER A NEGATIVE QUESTION

This paper deals with answers to negative *yes-no* questions, focusing on data from Czech. It is shown that answering particles can express both positive and negative answers to negative questions, but that their choice is not free. Several pieces of evidence are discussed in order to show that the use of the particles depends on the interpretation of negation in the question: expletive negation or true negation. This semantic distinction is furthermore tightly linked to the syntactic position of the negation, according to which we distinguish between negative interrogative clauses and negative declarative clauses used as questions. An analysis in terms of absolute and relative polarity is proposed to account for the mixed behaviour of answering particles: particles express absolute polarity in answers to interrogative questions, whose polarity is open, and relative polarity in answers to declarative questions, whose polarity has been already specified.

**Keywords:** yes-no question, answer, answering particles, negation, polarity

### Povzetek

#### DA ALI NE, ALI KAKO ODGOVORITI NA NIKALNO VPRAŠANJE

Pričujoči članek obravnava odgovore na nikalna *da-/ne*-vprašanja s posebnim poudarkom na češčini. Obravnava primerov pokaže, da členka *ano* in *ne*, ki se uporabljata v odgovorih na nikalna *da-/ne*-vprašanja, lahko izražata tako soglašanje kot zavračanje, vendar njuna izbira ni poljubna. Podatki pokažejo, da je raba posameznega členka odvisna od vrste zanikanja v vprašanju: to je lahko pravo/stavčno zanikanje ali pleonastično zanikanje. Različna pomenska interpretacija členka je odvisna tudi od skladenjskega položaja nikalnice. Glede na ta kriterij avtorica loči nikalne vprašalne povedi od nikalnih povednih povedi, ki se uporabljajo kot vprašanja. Različno rabo in interpretacijo členkov avtorica pojasni z vidika absolutne in relativne polarnosti: v odgovorih na povedi z vprašalnico, katerih polarnost je odprta, členka izražata absolutno polarnost, medtem ko izražata relativno polarnost v odgovorih na povedna vprašanja, ki so rabljena vprašalno in katerih polarnost je že predhodno specificirana.

**Ključne besede:** *da-/ne*-vprašanja, odgovori, členki, nikalnost, polarnost



## TENSE AND MODALITY IN THE NOMINAL DOMAIN\*\*

### 1. INTRODUCTION

Tense and modality have been topics of high interest to linguists. The semantic and syntactic properties of these elements and their interaction at the sentential level have been widely investigated (on tense, see, among many others, Stowell 1982, 1992, 1995, 2000, 2007; Comrie 1985; Abusch 1988, 1997; Ogihara 1995; Steedman 1997; and on modality, see Kratzer 1977, 1981, 1991, 2012, 2013; Palmer 1990, 2001; Nuyts 1993, 2001, 2006; Narrog 2005; von Stechow 2006; Hacquard 2006, 2010, 2011; Portner 2009; among others). Crucially, it has been convincingly argued that in the cartography of syntactic structure, epistemic modals are base-generated and interpreted higher than Tense (T), whereas root (non-epistemic) modals are base-generated and interpreted lower than T (e.g., Cinque 1999, 2004, 2013; Butler 2003; Hacquard 2010).

At the subsentential level, however, the literature is limited to the study of temporality and mood in the nominal domain and their syntactic representations as nominal tense and nominal mood, leaving much space for research and discoveries. Musan (1995, 1997, 1999) observes that noun phrases can be modified by various kinds of temporal expressions and that these temporal modifiers can modify nouns denoting life-time or temporary properties. Lecarme (1996, 2004, 2008), Sadler and Nordlinger (2001) and Nordlinger and Sadler (2003, 2004a, 2004b) point out that nominals are inflected for tense, aspect and mood in a number of languages (e.g., Halkomelem and Guaraní). Contra this, Alexiadou (2001, 2005), based on the absence of phenomena related to T (namely, Extended Projection Principle (EPP), Exceptional Case Marking (ECM) and raising) in nominals, argues that these constructions lack a tense projection.

The main objective of this article is to show that, similar to the structure of Complementizer Phrases (CPs), epistemic and root modal elements have different positions in Determiner Phrases (DPs); epistemic adjectives appear in the specifier of  $\text{Mod}_{\text{epis.N}}^{\text{P}}$  above nominal tense ( $\text{T}_N\text{P}$ ), while root adjectives appear in the specifier of  $\text{Mod}_{\text{root.N}}^{\text{P}}$  below  $\text{T}_N\text{P}$ . With this aim in view, the structure of the article is as follows. In section 2, we will have a brief look at tense and modality in the verbal domain to which I seek some parallelism in the nominal domain. In section 3, I elaborate on the concept of

---

\* n.ilkhani pour@ut.ac.ir

\*\* A previous version of this article was presented at SinFonIJA 8 in Ljubljana, 24-26 September 2015. I would like to thank the organizers and the audience of that conference for their feedback. Also, I am grateful to two anonymous reviewers of *Linguistica* for their insightful comments.



temporality in nominals. I importantly discuss that we need some extending-into-time perspective in noun phrases and that this extending-into-time perspective is syntactically represented as nominal tense. In section 4, I move on to some data from Persian and show that the temporal adjective *qæbli* ‘previous’ is, in effect, ambiguous. In section 5, based on the ambiguity of the temporal adjective *qæbli* ‘previous’, I argue that epistemic and root adjectives (e.g., *Zehtemali* ‘probable’ and *qabel-e-ZeZtemad* ‘reliable’, respectively) occupy different positions with regard to the nominal tense. Finally, in section 6, I wrap up the article with some concluding remarks.

## 2. THEORETICAL PRELUDE

The literature on tense and modality in the verbal domain is vast and burgeoning. In this section, I hardly attempt to summarize the diverse and insightful findings of previous studies. Rather, I wish to pursue the idea that epistemic modals ( $\text{Mod}_{\text{epis}}$ ) are base-generated higher than root modals ( $\text{Mod}_{\text{root}}$ ) in the clausal structure. To put it precisely, as illustrated in (1), epistemic modals are higher than T, while root modals are lower than T (e.g., Cinque 1999, 2004, 2013; Butler 2003; Portner 2009).<sup>1</sup>

$$(1) \text{Mod}_{\text{epis}} < T < \text{Mod}_{\text{root}}$$

Maintaining Kratzer’s (1981, 1991) original account that each modal has a single lexical entry which is not specified for a particular flavor (epistemic or root), Hacquard (2010) shows that modal auxiliaries, such as *must* and *can*, may freely appear above or below T and that the individual relativity and the time relativity of modals go hand in hand, yielding the event relativity of modal elements: epistemic modals, base-generated in the high position above TP, are evaluated in the context of the speech event, that is, with regard to the speaker at the speech time, while root modals, base-generated in the low position above VP, are evaluated in the context of the VP event, with regard to one of the participants of the event, represented as an argument, at the event time. This is illustrated in (2).

$$(2) \underbrace{[\text{AssertP} [\text{Mod}_{\text{epis}}\text{P} [\text{TP} [\text{Mod}_{\text{root}}\text{P} [\text{VP} \dots ]]]]]]}_{\text{Speech event (speaker, speech time)}} \quad \underbrace{[\text{TP} [\text{Mod}_{\text{root}}\text{P} [\text{VP} \dots ]]]]}_{\text{VP event (event participant = argument, event time)}}$$

According to Hacquard’s (2010) analysis, the ambiguity of (3) for the modal readings (epistemic and root) is due to the two different structural positions that *must* may occupy in the derivation: when merged above T, it is evaluated in the context of the speech event, as in (3a), and when merged lower than T, it is evaluated with regard to the VP event, as in (3b).

<sup>1</sup> I remain agnostic to the finer distinctions of epistemic and root modals (see Cinque 1999, 2004, 2013).

(3) *The winners must have been in the court at noon.*

- a.  $\text{Mod}_{\text{epis}} < \text{T}$ : epistemic reading: it is necessary, given what is known now by the speaker, that the winners were in the court at noon.
- b.  $\text{T} < \text{Mod}_{\text{root}}$ : root reading: it was necessary, given the rules of the event then, that the winners were in the court at noon.

In the rest of this article, I argue that parallel to the structure of the verbal domain in (2), epistemic and root modal elements occupy different positions in DPs; epistemic adjectives appear above nominal tense and are interpreted in the context of the speech event, while root adjectives appear in a lower position, interpreted in the context of the NP event. This is illustrated in (4) where the subscript N indicates the nominal stance of the projection.



To reach (4), first we should ascertain that temporality is an indispensable concept in noun phrases and that the temporal dimension of nominals is syntactically represented via Nominal Tense ( $\text{T}_{\text{N}}$ ).

### 3. TEMPORALITY AND TENSE IN NOMINALS

The modification of noun phrases with temporal expressions and the temporal relation between predicates and their arguments are two pieces of evidence for considering an extending-into-time perspective for nominals.

First, as observed by Musan (1995), noun phrases can be modified by various kinds of temporal expressions, as shown in (5). These temporal modifiers can modify nouns denoting life-time or temporary properties, as in (5a) and (5b), respectively.

- (5) a) clausal modifiers: [*The war when my grandfather was young lasted four years.*
  - b) genitive modifiers: [*The sixties' rebels are quite established today.*
  - c) adverbial modifiers: [*The quarrel yesterday was totally superfluous.*
  - d) adjectival modifiers: [*The present wife of Klaus] is [a former student of his].*
  - e) prepositional modifiers: [*The chancellor in 1989 made some serious mistakes.*
- (Musan 1995:160, (1a–e))

The temporal modification of noun phrases leaves space for the hypothesis that nominals, similar to verbals, refer to situations that hold at certain times (see also Musan 1999).

Second, the unacceptability of the sentences in (6) can be justified if we credit the arguments *my future job* and *the present president* with some temporal dimension modified by the temporal adjectives *future* and *present* (see also Enç 1987).

- (6) a) #*Last week I was dismissed from my future job.*  
 b) #*The present president will be elected (president) shortly.* (where every person can be president only once)

A relevant discussion of the temporal relation between predicates and their arguments can be found in Musan (1995, 1997, 1999) where EXISTENCE-IMPLYING predicates, including individual-level predicates (e.g., *to be from America*) and stage-level predicates (e.g., *to be happy*), are assumed to impose a presuppositional condition on their arguments' being in existence or alive. Musan also points out that noun phrases can receive temporally dependent or independent readings:

A noun phrase occurrence is temporally dependent if and only if its situation time has to intersect with the situation time of the main predicate of its clause.

A noun phrase occurrence is temporally independent if and only if its situation time does not have to intersect with the situation time of the main predicate of its clause. (Musan 1999:622)

Now the question is whether there exists a tense projection in the nominal spine. Alexiadou (2001, 2005) points out that there are two features associated with T: the EPP feature and the [assign nominative] feature. She further discusses that the phenomena related to T, namely EPP, ECM and raising, are absent in nominals and thus noun phrases lack a tense projection (Alexiadou 2001:59–66).

- (7) EPP in CP and DP

- a) *There arrived a man.*  
 b) \**there's arrival*

(Alexiadou 2001: 60, (83a–b))

- (8) ECM in CP and DP

- a) *I believe Mickey to be a genius.*  
 b) \**my belief of Mickey to be genius*

(Alexiadou 2001: 60, (88) and (89))

- (9) Raising in CP and DP

- a) *Mary appears to have left.*  
 b) \**Mary's appearance to have left*

(Alexiadou 2001: 60, (84) and (85))

While Alexiadou (2001) takes these as pieces of evidence for the absence of tense in noun phrases, I take them as evidence for some featural difference between clausal and nominal tense. Recall that D is the nominal counterpart of C and that D and C do not share the same features or interpretations. So, I submit that it is possible that  $T_N$  and T do not share the same features or interpretations and consequently, that  $T_N$  is the nominal counterpart of T.

On the other hand, the presence of tense in the extended projection of noun phrases is fruitful. Looking with favour upon Cinque's (1994, 2010) and Scott's (2002) proposal that adjectives are base-generated in the specifiers of distinct functional projections to which they are associated, I consider that the location of temporal adjectives (e.g., *present* and *former* in (5d)) is Spec,T<sub>N</sub>P. Nominal Tense is then a functional projection that not only provides the necessary space for temporal adjectives but also contributes to the temporal interpretation of nouns.

Moreover, according to Lecarme (1996, 2004, 2008), Sadler and Nordlinger (2001) and Nordlinger and Sadler (2003, 2004a, 2004b), among others, nominals are inflected for tense, aspect and mood in a number of languages. In Halkomelem and Guaraní, as shown in (10) and (11), the same set of affixes mark tense on nominals and verbal predicates. Nominal past tense encoding meanings such as 'former, -ex, late (dead)' temporally locates the nominal. When used with a possessed inanimate noun, the temporal marker indicates that the possession relation was in the past, or that the possessed item has been destroyed (Burton 1997:67–68).<sup>2</sup>

- (10) *te sqwemá:y-elh* (Halkomelem)  
 the dog-pst  
 'the dead dog'  
 (Sadler and Nordlinger 2001)

- (11) *che-róga-kue* (Guaraní)  
 1sg-house-pst  
 'my former house'  
 (from Nordlinger and Sadler 2004a)

The morphological expression of time within the nominal domain provides evidence for the idea that T does *not* exclusively combine with verbs (cf. Panagiotidis 2011, 2015).

And last but not least, the study of temporality in the nominal domain and assuming that T<sub>N</sub> is a functional projection in the nominal spine provide an exciting opportunity to advance our knowledge of the possible orders of adjectives. In particular, we find the way to realize that epistemic adjectives are higher than T<sub>N</sub>, whereas root adjectives are lower than T<sub>N</sub>. To achieve this order, I start the argument with discussing the ambiguity of the Persian adjective *qæbli* 'previous' in the next section.

#### 4. PERSIAN 'PREVIOUS': ORDINAL/TEMPORAL ADJECTIVE

The Persian adjective *qæbli* 'previous' in (12) is ambiguous; it may have a temporal reading, as in context A, or an ordinal reading, as in context B.

2 Tonhauser (2006, 2007) argues that nominal temporal markers in Paraguayan Guaraní are aspect (and not tense) markers. It is beyond the scope of this study to decide on aspect/tense analysis of nominal temporal markers in this language.

- (12) *bærænde-ye qæbli*  
 winner-EZ previous  
 ‘the previous winner’

Context A: The Temporal Reading

The host of a quiz show talks about the winner of the previous round.

In this context, the host says the sentence in (13).

- (13) *bærænde-ye qæbli tehrani bud*  
 winner-EZ previous Tehrani be.PST.3SG  
 ‘The previous winner was from Tehran.’

In context A, being the winner is referred to a property in the past. The previous winner does not hold the title any more in the speech time of (13).

Context B: The Ordinal Reading

In a quiz show with two winners, the host introduces first winner 1 and then winner 2.

While introducing winner 2 in context B, the host says the sentence in (14).

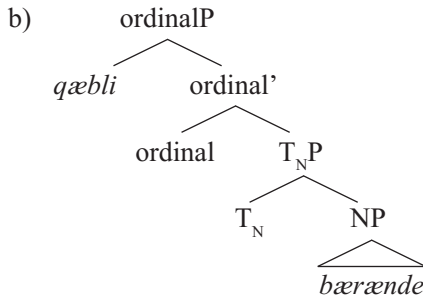
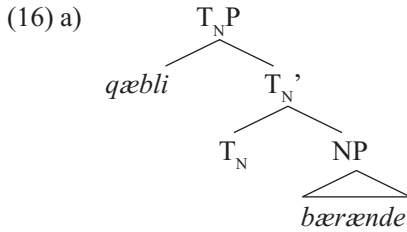
- (14) *Ûin bærænde mesl-e bærænde-ye qæbli tehrani-e*  
 this winner like-EZ winner-EZ previous Tehrani-be.3SG  
 ‘This winner, like the previous winner, is from Tehran.’

Here, the property of being the winner in the DP *bærænde-ye qæbli* ‘the previous winner’ is held at the speech time. The adjective *qæbli* ‘previous’ does not make being the winner a property in the past. Instead, it puts an order on the two winners and thus, plays the role of an ordinal number. The sentence in (14) as put in context B can be paraphrased as in (15).

- (15) *Ûin bærænde mesl-e bærænde-ye Ûævvæl tehrani-e*  
 This winner like-EZ winner-EZ first Tehrani-be.3SG  
 ‘This winner, like the first winner, is from Tehran.’

The ambiguity of (12) does not come from two distinct lexical entries for the adjective *qæbli* ‘previous’. It is indeed due to the two different positions that this adjective may occupy in the course of the syntactic derivation of the DP: the specifier of  $T_NP$  in context A, as in (16a), that gives rise to the temporal reading, and the specifier of ordinalP in context B, as in (16b), that brings about the ordinal reading.<sup>3</sup>

3 In this article, I disregard the syntactic nature of *Ezafê*. For discussions on this element in Persian, see Samiiian 1994, Ghomeshi 1997, Abolghasemi 2002, Larson and Yamakido 2005, DeLazero and Geraee 2014, and references therein. However, it is worth mentioning that the occurrence of *Ezafê* independent of the syntactic and semantic nature of the elements in the nominal domain (consider the fact that *Ezafê* appears on nouns and also on adjectives, and that it appears on both root and epistemic modal adjectives, as well as temporal adjectives) assures that this element does not play any role in the discussion of this research.



The proposal that the adjective *qæbli* ‘previous’ can be base-generated in two structural positions is supported by the co-occurrence of the two readings.<sup>4</sup> Consider context C.

Context C: Co-Occurrence of the Temporal and Ordinal Readings

In a quiz show, the host introduces the two winners of the previous round, first winner 1 of the previous round and then winner 2 of the previous round.

In this context, the host refers to winner 2 of the previous round as in (17).

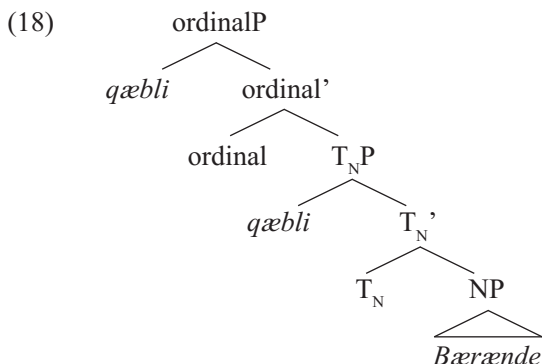
- (17) *bærænde-ye qæbli-ye qæbli*  
 winner-EZ previous-EZ previous  
 ‘the previous previous winner’

Here, the first *qæbli* ‘previous’ (the one closer to the modified noun) indicates that the property of being the winner was held in the past (and not at the speech time) and the second *qæbli* ‘previous’ puts an order on the two winners of the previous round. The syntactic positions of the two adjectives are illustrated in (18).<sup>5</sup>

4 I am thankful to Marcel den Dikken for bringing this point to my attention, and to Guglielmo Cinque for a short but thoughtful discussion on this.

5 A reviewer pointed out that the ordinal *previous* and the temporal *previous* can be separated by a numeral in English: *the previous two previous winners* (see also Cinque 2015:24, fn. 4). This cannot be observed in Persian because numerals precede nouns, while temporal and ordinal adjectives follow them. The order in Persian is as illustrated in (i) and exemplified in (ii) below. The example in (ii) is ambiguous: it can be read as in (ii.a), or as in (ii.b).

(i) *Number Noun-EZ Adjective-EZ Adjective*



In the next section, I will employ the structural ambiguity of *qæbli* ‘previous’ to show that epistemic adjectives are base-generated higher than  $T_N$  while root adjectives occupy a lower position.<sup>6</sup>

## 5. MODAL ADJECTIVES IN ORDER

In Persian, the notion of modality can be expressed via modal adjectives in the nominal domain. Examples (19a–b) are instances of epistemic and root modal adjectives in Persian (Ilkhanipour and DeLazero 2013; Ilkhanipour 2015).

- (19) a) *bærænde-ye ʔehtemali*  
 winner-EZ probable  
 ‘the probable winner’
- b) *bærænde-ye qabel-e-ʔeʔtemad-e*  
 winner-EZ reliable  
 ‘the reliable winner’

In the following lines of this section I argue for the fixed order of adjectives in (20), and since I regard adjectives as merged in the specifiers of functional projections, for the hierarchy of functional projections in the nominal spine in (21).

- 
- (ii) *do bærænde-ye qæbli-ye qæbli*  
 two winner-EZ previous-EZ previous

a. *do [bærænde-ye qæbli-ye qæbli]*  
 ‘the two previous previous winner’

b. *[do bærænde-ye qæbli]-ye qæbli*  
 ‘the previous two previous winner’

6 It should be noted that the word order in Persian noun phrases (e.g., *Noun-EZ Adjective-EZ Adjective* in (17)) does not conflict with the “structural” hierarchy of the syntactic elements (e.g., as illustrated in (18)). My assumption here is “[...] that what reaches the mind lacks order, while what reaches the ear is ordered. Linear order, then, should not enter into the syntactic-semantic computation. Rather, it is imposed by externalization, presumably as a reflex of properties of the SM system, which requires linearization [...]” (Chomsky 2015:19).

(20) ordinal adjective < epistemic adjective < temporal adjective < root adjective

(21) ordinalP < Mod<sub>epis.N</sub>P < T<sub>N</sub>P < Mod<sub>root.N</sub>P

### 5.1. Root Adjectives: Below T<sub>N</sub>

The structural ambiguity of *qæbli* ‘previous’ is observed when it co-occurs with root adjectives, such as *qabel-e-ʔe ʔtemad* ‘reliable’, as in (22).

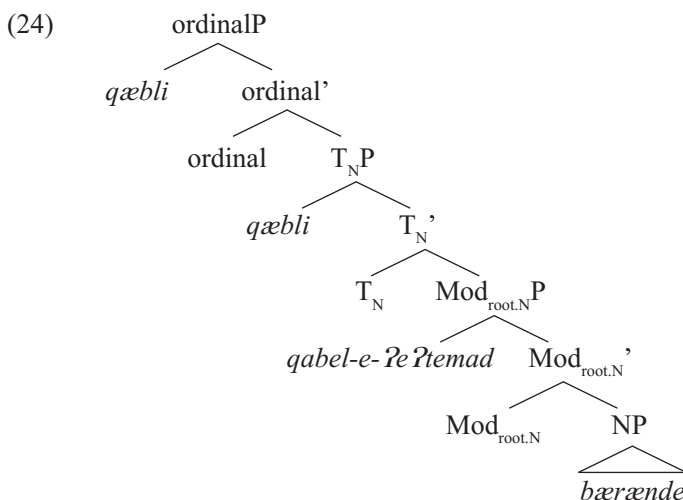
(22) *bærænde-ye qabel-e-ʔe ʔtemad-e qæbli*  
 winner-EZ reliable-EZ previous  
 ‘the previous reliable winner’

The DP in (22) can be uttered in both contexts A and B, as defined above. Importantly, it is acceptable in context A with the temporal reading of *qæbli* ‘previous’.

Moreover, the root adjective *qabel-e-ʔe ʔtemad* ‘reliable’ can modify the noun in context C where the two readings of *qæbli* ‘previous’ co-occur. This can be seen in (23).

(23) *bærænde-ye qabel-e-ʔe ʔtemad-e qæbli-ye qæbli*  
 winner-EZ reliable-EZ previous-EZ previous  
 ‘the previous previous reliable winner’

These pieces of evidence suggest that the position of the root adjective is lower than T<sub>N</sub>, as shown in (24).



The root adjective *qabel-e-ʔe ʔtemad* ‘reliable’ is then interpreted with respect to the modified noun *bærænde* ‘winner’ at the time of the quiz show. The quiz show is, in effect, the NP event against which the root adjective is evaluated. The individual



involved in this event is the role noun *bærænde* ‘winner’ and the time of the event is syntactically represented as  $T_N$ , the time of being the winner or winning here.

Reordering the temporal and the root adjectives, as in (25), we still find both readings of *qæbli* ‘previous’ available. The phrase in (25) is marked, though, with heavy stress on the root adjective *qabel-e-ʔe ʔtemad* ‘reliable’.

- (25) *?bærænde-ye qæbli-ye qabel-e-ʔe ʔtemad*  
 winner-EZ previous-EZ reliable  
 ‘?the reliable previous winner’

The order of adjectives in (25) can be accounted for if we consider a focus phrase in the nominal domain and move the root adjective from its original position to the specifier of the focus phrase, as illustrated in (26) (on focus in nominals, see Aboh 2004; Giusti 2005; Angitso 2015; among others).<sup>7</sup>

<sup>7</sup> As correctly pointed out by a reviewer, what remains to be determined is whether this focus position is available as a means to subvert the order of direct modification adjectives. To check that, one should select two exclusively non-predicative adjectives, otherwise it could simply be that the apparent order subversion is a function of using one as a direct modification and one as an indirect (reduced relative clause) modification (cf. Cinque 2010:32). In Persian temporal and ordinal adjectives are the only instances of exclusively non-predicative adjectives. As can be seen in (i) and (ii), *qæbli* ‘previous’ and *bæʔdi* ‘next’ with both their temporal and ordinal readings cannot occur in the predicate position.

- (i) a) *modir-e qæbli*  
 manager-EZ previous  
 ‘the previous manager’  
 b) *\*modir qæbli ʔæst*  
 manager previous be.3SG  
 ‘\*The manager is previous.’  
 (ii) a) *modir-e bæʔdi*  
 manager-EZ next  
 ‘the next manager’  
 b) *\*modir bæʔdi ʔæst*  
 manager next be.3SG  
 ‘\*The manager is next.’

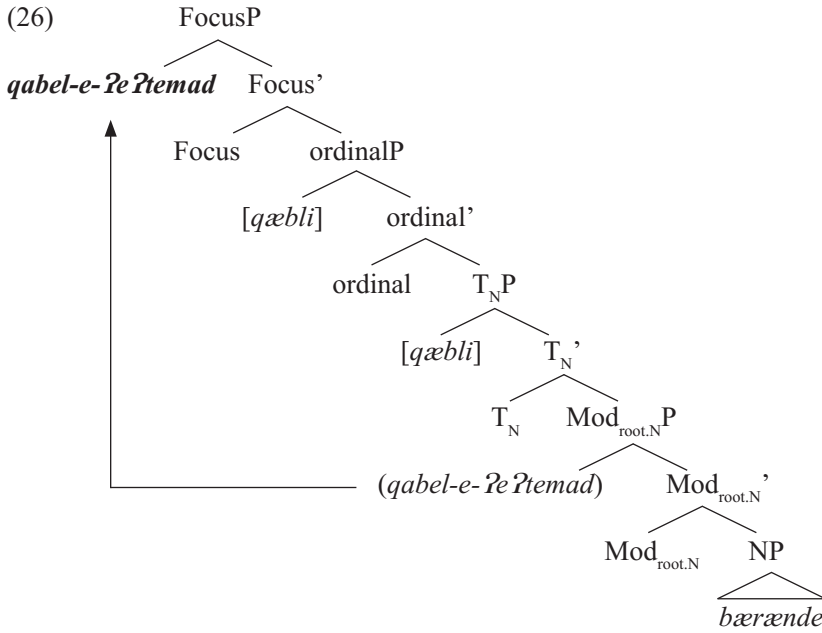
Now consider employing *qæbli* ‘previous’ with its temporal reading and *bæʔdi* ‘next’ with its ordinal reading co-modifying the noun *modir* ‘manager’ in a context where there exists two previous managers and the speaker introducing the second previous manager utters the DP in (iii).

- (iii) *modir-e qæbli-e bæʔdi*  
 manager-EZ previous-EZ next  
 ‘the next previous manager’

The temporal adjective *qæbli* ‘previous’ in (iii) can be focalized; it moves from Spec, $T_N$ P to the Spec,FocusP and is pronounced with heavy stress, as illustrated in (iv).

- (iv) *?modir-e bæʔdi-e qæbli*  
 manager-EZ next-EZ previous

So, as can be seen, FocusP can be used as a means to subvert the order of direct modification adjectives in Persian.



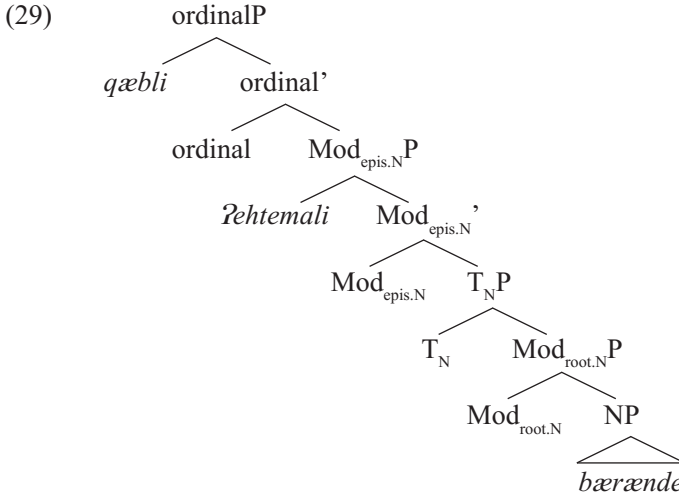
## 5.2. Epistemic Adjectives: Above $T_N$

With epistemic adjectives, such as *ʔehtemali* ‘probable’, as in (27), the adjective *qæbli* ‘previous’ is not ambiguous; it can be interpreted only as in context B, that is, as an ordinal modifier.

- (27) *bærænde-ye ʔehtemali-ye qæbli*  
 winner-EZ probable-EZ previous  
 ‘the previous probable winner’

The unacceptability of (27) in context A where the adjective *qæbli* ‘previous’ is interpreted as a temporal modifier and the fact that the epistemic adjective *ʔehtemali* ‘probable’ cannot be used in context C where the two readings of *qæbli* ‘previous’ co-occur, as shown in (28), suggest that the position of epistemic adjectives is higher than  $T_N$ , but lower than ordinalP. This is illustrated in (29).

- (28) *#bærænde-ye ʔehtemali-ye qæbli-ye qæbli*  
 winner-EZ probable-EZ previous-EZ previous  
 ‘#the previous previous probable winner’



Base-generated higher than  $T_N$ , the epistemic adjective is evaluated with regard to the speaker's knowledge at the speech time.

Interestingly, as shown in (30), epistemic adjectives cannot be focalized crossing over ordinalP; the adjective *qæbli* 'previous' contains a presuppositional condition on the modified noun and is EXISTENCE-IMPLYING (see Musan 1997, 1999): what already exists cannot be probable.

- (30) #*bærænde-ye qæbli-ye Zehtemali*  
 winner-EZ previous-EZ probable  
 intended: 'the previous probable winner'

### 5.3. Persian 'Possible': Epistemic/Root Adjective

As discussed in Hacquard (2010), considering different structural positions for epistemic and root modal auxiliaries (e.g., *can* and *must*) in the verbal spine disambiguates the epistemic/root meaning of these elements. In a similar vein, considering different structural positions for epistemic and root modal adjectives in the nominal spine disambiguates the epistemic/root meaning of the Persian modal adjective *momken* 'possible'. This adjective can be interpreted as an epistemic or a root modal adjective, as shown in (31).

- (31) *pasox-ha-ye momken dær Zin Zemtahan*  
 answer-PL-EZ possible in this exam  
 'the possible answers in this exam'

- a) epistemic reading: the answers that are possible, given what is known by the speaker at the speech time.
- b) root reading: the answers that are possible, given the rules/circumstances of the examination event at the exam time.

The analysis presented in this article accounts for the two readings of the adjective *momken* ‘possible’ in (31); when merged in  $\text{Spec,Mod}_{\text{epis.N}}\text{P}$ , it is interpreted as in (31a), and when merged in  $\text{Spec,Mod}_{\text{root.N}}\text{P}$ , it is interpreted as in (31b). Thus, we can say that in the lexicon there is only one *momken* which is neutral with respect to the modal base, and that the epistemic and the root meanings of this adjective come from the two different syntactic positions it may occupy in the functional hierarchy of the noun phrase.

## 6. CONCLUSION

In this article, I showed that the modification of noun phrases with temporal expressions and the temporal relation between predicates and their arguments provide evidence for considering some extending-into-time perspective in the nominal domain. Then, on the basis of the location-in-specifier approach to adjectival syntax and the morphological expression of time within the nominal domain, I indicated that tense combines with nominals although  $T_N$  does not share similar features with its clausal counterpart. After that, I discussed that the adjective *qæbli* ‘previous’ is structurally ambiguous: in the specifier of  $T_N\text{P}$ , it is interpreted as a temporal modifier and in the specifier of ordinalP, it is interpreted as an ordinal modifier. This structural ambiguity put forth a justification for the fixed order of the epistemic and root modal adjectives with regard to the temporal adjective. Root adjectives are base-generated lower than  $T_N$  in the specifier of  $\text{Mod}_{\text{root.N}}\text{P}$  and are evaluated with regard to the NP event; the time of the event is signaled by  $T_N$ , which refers to the existence/occurrence time of the modified noun, and the individual is the modified noun. Epistemic adjectives are base-generated higher than  $T_N$  in the specifier of  $\text{Mod}_{\text{epis.N}}\text{P}$  and are evaluated in the context of the speech event, that is, with regard to the speech time and the speaker’s knowledge. This is what I proposed in (4), repeated here in (32).

$$(32) \quad \underbrace{[\text{Assert}_N\text{P} [\text{Mod}_{\text{epis.N}}\text{P} [T_N\text{P} [\text{Mod}_{\text{root.N}}\text{P} [NP \dots ]]]]]}_{\text{Speech event}} \quad \underbrace{[T_N\text{P} [\text{Mod}_{\text{root.N}}\text{P} [NP \dots ]]]}_{\text{NP event}}$$

The last remark I wish to make here is that the nominal assertion  $\text{Assert}_N$  is a logical operator, co-indexed with its clausal counterpart, *Assert* (see Hacquard’s 2010). The question that may arise is how a noun phrase can be asserted, or more precisely, how a speech act mood projection can be present in the nominal domain (if at all). I leave the semantic nature of  $\text{Assert}_N$  and its relation to D for further research.

## References

- ABOH, Enoch (2004) “Topic and Focus within D.” *Linguistics in the Netherlands* 21, 1–12.  
 ABOLGHASEMI, Mohsen (2002) *tarix-e moxtæsxær-e zæban-e farsi* [The brief history of Persian]. Tehran: Tahoori.

- ABUSCH, Dorit (1988) "Sequence of tense, intensionality, and scope." *Proceedings of the West Coast Conference on Formal Linguistics* 7, 1–14.
- ABUSCH, Dorit (1997) "Sequence of tense and temporal de re." *Linguistics and Philosophy* 20, 1–50.
- ALEXIADOU, Artemis (2001) *Functional structure in nominals: Nominalization and ergativity*. Amsterdam: John Benjamins Publishing Company.
- ALEXIADOU, Artemis (2005) *Tense (and mood) in the nominal domain(?)*. Paper presented at The Workshop on the Linguistic Representation of Tense and Mood, Stuttgart.
- ANGITSO, Michael T. (2015) "The limits of the DP/CP parallelism: Evidence from Tiv." *Semantics-Syntax Interface* 2/2, 141–156.
- BURTON, Strang C. (1997) "Past tense on nouns as death, destruction, and loss." In: *Proceedings of NELS 27*. Amherst, MA: University of Massachusetts, 65–77.
- BUTLER, Jonny (2003) "A minimalist treatment of modality." *Lingua* 113, 967–996.
- CINQUE, Guglielmo (1994) "Evidence for partial N-movement in the Romance DP." In: G. Cinque/J. Koster/J. Pollock/L. Rizzi/R. Zanuttini (eds), *Paths towards universal grammar: Studies in honor of Richard S. Kayne*. Washington, D.C.: Georgetown University Press, 85–110.
- CINQUE, Guglielmo (1999) *Adverbs and functional heads: A crosslinguistic perspective*. New York and Oxford: Oxford University Press.
- CINQUE, Guglielmo (2004) "Issues in adverbial syntax." *Lingua* 114, 683–710.
- CINQUE, Guglielmo (2010) *The syntax of adjectives: A comparative study*. Cambridge, MA: MIT Press.
- CINQUE, Guglielmo (2013) "Cognition, universal grammar, and typological generalizations." *Lingua* 130, 50–65.
- CINQUE, Guglielmo (2015) "A note on 'other'." In: E. Brandner/A. Czipionka/C. Freitag/ A. Trotzke (eds), *Charting the Landscape of Linguistics: Webschrift for Josef Bayer*. University of Konstanz, 22–27.
- COMRIE, Bernard (1985) *Tense*. Cambridge: Cambridge University Press.
- DELAZERO, Octav E./Parvane GERAEE (2014) "Ezafé constructions as relative clauses: Semantic arguments." *Semantics-Syntax Interface* 1/1, 75–81.
- ENÇ, Mürvet (1987) "Anchoring conditions for tense." *Linguistic Inquiry* 18, 633–657.
- VON FINTEL, Kai (2006). "Modality and language." In: D.M. Borchert (ed.), *Encyclopedia of Philosophy (2nd ed.)*. Detroit: MacMillan Reference USA.
- GHOMESHI, Jila (1997) "Non-projecting nouns and the Ezafé construction in Persian." *Natural Language and Linguistic Theory* 15, 729–788.
- GIUSTI, Giuliana (2005) "At the left periphery of the Romanian noun phrase." In: M. Coene/L. Tasmowski (eds), *On Space and Time in Language*. Cluj: Clusium, 23–49.
- HACQUARD, Valentine (2006) *Aspects of modality*. PhD dissertation, MIT, Cambridge, MA.
- HACQUARD, Valentine (2010) "On the event relativity of modal auxiliaries." *Natural Language Semantics* 18/1, 79–114.

- HACQUARD, Valentine (2011) “Modality.” In: C. Maienborn/K. von Stechow/P. Portner (eds), *Semantics: An international handbook of natural language meaning*. Berlin: Mouton de Gruyter, 1484–1515.
- ILKHAPOUR, Negin (2015) *sefat-e væjhi dær zæban-e farsi* [Modal adjectives in Persian]. Tehran: Nashr-e-Markaz.
- ILKHAPOUR, Negin/Octav E. DELAZERO (2013) *Selectional properties of Persian modal adjectives*. Talk presented at ICIL5, Bamberg.
- KRATZER, Angelika (1977) “What ‘must’ and ‘can’ must and can mean.” *Linguistics and Philosophy* 1, 337–355.
- KRATZER, Angelika (1981) “The notional category of modality.” In: H.-J. Eikmeyer/H. Rieser (eds), *Words, worlds, and contexts: New approaches in word semantics*. Berlin: Walter de Gruyter, 38–74.
- KRATZER, Angelika (1991) “Modality.” In: A. von Stechow and D. Wunderlich (eds), *Semantics: An international handbook of contemporary research*. Berlin: de Gruyter, 639–650.
- KRATZER, Angelika (2012) *Modals and conditionals: New and revised perspectives*. Oxford: Oxford University Press.
- KRATZER, Angelika (2013) *Modality for the 21st century*. Keynote talk at the 19th International Congress of Linguistics, Geneva.
- LARSON, Richard/Hiroko YAMAKIDO (2005) *Ezafé and the deep position of nominal modifiers*. Barcelona Workshop on Adjectives and Adverbs.
- LECARME, Jacqueline (1996) “Tense in the nominal system: The Somali DP.” In: J. Lecarme/J. Lowenstamm/U. Shlonsky (eds), *Studies in Afroasiatic grammar*. The Hague: Holland Academic Graphics, 159–178.
- LECARME, Jacqueline (2004) “Tense in nominals.” In: J. Guéron and J. Lecarme (eds), *The syntax of time*. Cambridge, MA: MIT Press, 440–475.
- LECARME, Jacqueline (2008) “Tense and modality in nominals.” In: J. Guéron/J. Lecarme (eds), *Time and modality*. Dordrecht: Springer, 195–225.
- MUSAN, Renate (1995) “On the temporal interpretation of noun phrases.” PhD dissertation, MIT.
- MUSAN, Renate (1997) “Tense, predicates, and lifetime effects.” *Natural Language Semantics* 5, 271–301.
- MUSAN, Renate (1999) “Temporal interpretation and information-status of noun phrases.” *Linguistics and Philosophy* 22, 621–661.
- NARROG, Heiko (2005) “On defining modality again.” *Language Sciences* 27, 165–192.
- NORDLINGER, Rachel/Louisa SADLER (2003) “The syntax and semantics of tensed nominals.” In: *Proceedings of LFG03 Conference*. State University of New York at Albany.
- NORDLINGER, Rachel/Louisa SADLER (2004a) “Nominal tense in cross-linguistic perspective.” *Language* 80/4, 776–806.
- NORDLINGER, Rachel/Louisa SADLER (2004b) “Tense beyond the verb: Encoding clausal tense/aspect/mood on nominal dependents.” *Natural Language and Linguistic Theory* 22, 597–641.

- NUYTS, Jan (1993) “Epistemic modal adverbs and adjectives and the layered representation of conceptual and linguistic structure.” *Linguistics* 31, 933–969.
- NUYTS, Jan (2001) *Epistemic modality, language, and conceptualization*. Amsterdam & Philadelphia: John Benjamins.
- NUYTS, Jan (2006) “Modality: Overview and linguistic issues.” In: W. Frawley (ed.), *The expression of modality*. Berlin: Mouton de Gruyter, 1–26.
- OGIHARA, Toshiyuki (1995) “The semantics of tense in embedded clauses.” *Linguistic Inquiry* 26, 663–679.
- PALMER, Frank R. (1990) *Modality and the English modals*. New York: Routledge.
- PALMER, Frank R. (2001) *Mood and modality (2nd ed.)*. Cambridge: Cambridge University Press.
- PANAGIOTIDIS, Phoivos (2011) “Categorial features and categorizers.” *The Linguistic Review* 28/3, 365–386.
- PANAGIOTIDIS, Phoivos (2015) *Categorial features: A generative theory of word class categories*. Cambridge: Cambridge University Press.
- PORTNER, Paul (2009) *Modality*. Oxford: Oxford University Press.
- SADLER, Louisa/Rachel NORDLINGER (2001) “Nominal tense with nominal scope: A preliminary sketch.” In: *Proceedings of LFG01 Conference*. University of Hongkong, Hongkong.
- SAMIHAN, Vida (1994) “The Ezafe construction: Some implications for the theory of X-bar syntax.” In: Mehdi Marashi (ed.), *Persian studies in North America*. Bethesda, MD: Iranbooks, 17–41.
- SCOTT, Gary-John (2002) “Stacked adjectival modification and the structure of nominal phrases.” In: G. Cinque (ed.), *Functional structure in DP and IP: The cartography of syntactic structures*, vol. 1. Oxford: Oxford University Press, 91–120.
- STEEDMAN, Mark (1997) “Temporality.” In: J. van Benthem/A. ter Meulen (eds), *Handbook of logic and language*. Amsterdam: Elsevier.
- STOWELL, Tim (1982) “On tense of infinitives.” *Linguistic Inquiry* 13, 561–570.
- STOWELL, Tim (1992) “Aspects of tense theory.” *GLOW Newsletter* 28, 48–49.
- STOWELL, Tim (1995) “The phrase structure of tense.” In: J. Rooryck/L. Zaring (eds), *Phrase structure and the lexicon*. Dordrecht: Kluwer, 277–291.
- STOWELL, Tim (2000) *Tense and modals*. Presented at International Round Table ‘The Syntax of Tense and Aspect’, Université Paris III - Sorbonne Nouvelle, 5–18 November.
- STOWELL, Tim (2007) “The syntactic expression of tense.” *Lingua* 114, 437–463.
- TONHAUSER, Judith (2006) *The temporal semantics of noun phrases: Evidence from Guaraní*. PhD dissertation, Stanford University.
- TONHAUSER, Judith (2007) “What is nominal tense? A case study of Paraguayan Guaraní.” In: *Proceedings of Semantics of Under-represented Languages (SULA) III*. Amherst, MA: GLSA Publications.



Summary  
TENSE AND MODALITY IN THE NOMINAL DOMAIN

It is well discussed in the literature that epistemic modals ( $\text{Mod}_{\text{epis}}$ ) are base-generated higher than Tense (T), while non-epistemic/root modals ( $\text{Mod}_{\text{root}}$ ) are base-generated lower than T, and that high modals are evaluated in the context of the speech event (i.e., with regard to the speaker at the speech time), whereas low modals are evaluated in the context of the VP event (with regard to an argument at the event time).

In this study, looking with favour upon the presence of tense and modal functional projections in the nominal domain, and following the idea that adjectives are base-generated in the specifiers of distinct functional projections, I argue that, similar to the structure of CPs, epistemic and root modal elements have different positions in DPs; epistemic adjectives appear in the specifier of  $\text{Mod}_{\text{epis.N}}\text{P}$  above nominal tense ( $\text{T}_\text{N}\text{P}$ ), while root adjectives appear in the specifier of  $\text{Mod}_{\text{root.N}}\text{P}$  below  $\text{T}_\text{N}\text{P}$ , where nominal tense is the time of the existence or occurrence of the modified noun.

With this aim in view, first, I show that the ambiguity of the adjective *qæbli* ‘previous’ is due to the two positions this adjective can occupy: the specifier of  $\text{T}_\text{N}\text{P}$  and the specifier of ordinalP, where the adjective receives temporal and ordinal interpretations, respectively.

Next, I explain that this structural ambiguity is observed when *qæbli* ‘previous’ co-occurs with root adjectives, such as *qabel-e- Ze Ztemad* ‘reliable’. This suggests that the position of root adjectives is lower than  $\text{T}_\text{N}$ , where it is interpreted with respect to the modified noun at the event time.

With epistemic adjectives, such as *Zehtemali* ‘probable’, the adjective *qæbli* ‘previous’ is not ambiguous; it can be interpreted only as an ordinal modifier. This implies that the epistemic modal is higher than  $\text{T}_\text{N}$ , where it is evaluated with regard to the speaker’s knowledge at the speech time.

Thus, we see that the interaction of temporal and modal adjectives in DPs provides evidence for a structural hierarchy in the nominal domain parallel to its counterpart at the clausal level.

**Keywords:** adjectives, modality, nominals, Persian, tense

Povzetek  
ČAS IN MODALNOST V SAMOSTALNIŠKI DOMENI

Znanstvene razprave pogosto umeščajo epistemske modalne prvine ( $\text{Mod}_{\text{epis}}$ ) v bazično tvorjen položaj nad Tense (T), medtem ko so neepistemske oziroma korenske modalne prvine ( $\text{Mod}_{\text{root}}$ ) bazično tvorjene nižje od T. Višje umeščene modalne prvine so ovrednotene v kontekstu govornega dogodka (v odnosu do govorca v času govora), medtem ko so nižje umeščene modalne prvine ovrednotene v kontekstu dogodka znotraj glagolske zveze VP (v odnosu do argumenta v času dogodka).



V pričujoči študiji, ki sledi pristopom, ki dopuščajo časovne in modalne funkcionalne projekcije v samostalniški domeni in privzemajo, da so pridevniki bazično tvorjeni v določilih ločenih funkcionalnih projekcij, trdimo, da imajo – podobno, kot je to znano za zgradbo vezniške zveze CP – epistemske in korenske modalne prvine v določilniških zvezah (DP) različne položaje: epistemski pridevniki se pojavljajo v določilu  $\text{Mod}_{\text{epis},N}P$  nad samostalniškim časom ( $T_N P$ ), medtem ko se korenski pridevniki pojavljajo v določilu  $\text{Mod}_{\text{root},N}P$  pod zvezo  $T_N P$ , kjer samostalniški čas predstavlja čas obstoja ali čas pojavitve modificiranega samostalnika.

Tako najprej pokažemo, da dvoumnost pridevnika *qæbli* ‘prejšnji’ izhaja iz dveh skladenjskih položajev, ki jih lahko ta pridevnik zavzame: določilo zveze  $T_N P$  in določilo zveze ordinalP, kjer pridevnik prejme časovno in vrstilniško interpretacijo.

Nato razložimo, da do zgradbene dvoumnosti pride, ko se pridevnik *qæbli* ‘prejšnji’ pojavi s korenskimi pridevniki, kakršen je *qabel-e-ʔe ʔtemad* ‘zanesljiv’. Tako pokažemo, da je položaj korenskih pridevnikov nižje od  $T_N$ , kjer je interpretiran v odnosu do modificiranega samostalnika v času dogodka.

Z epistemskimi pridevniki, kakršen je *ʔehtemali* ‘verjeten’, pridevnik *qæbli* ‘prejšnji’ ni dvoumen; lahko je interpretiran le kot vrstilniški modifikator. To pokaže, da je epistemska prvina višje od  $T_N$ , kjer je ovrednotena v odnosu do govorčevega vedenja v času govora.

V prispevku tako pokažemo, da interakcija časovnih modalnih pridevnikov v zvezi DP dokazuje skladenjsko hierarhijo v samostalniški domeni, ki je vzporedna tisti na ravni stavka.

**Ključne besede:** pridevniki, modalnost, nominalne prvine, perzijščina, čas



## THE INTENSIONAL PROFILES OF FIVE HUNGARIAN IMPERATIVE SENTENCE TYPES<sup>1</sup>

### 1. INTRODUCTION

The purpose of the paper is to provide formal pragmatico-semantic analysis for five Hungarian imperative sentence types. Imperatives can be associated with a wide range of speech acts. Our aim is to capture and formalize the information states – beliefs, desires, and intentions – behind the baseline imperative, along with four further types modified by discourse markers or non-standard intonation. The *intensional profiles* which characterize these five types of Hungarian imperatives are described in the (S)DRT-based discourse-semantic framework of  $\Re$ ALIS.

In the paper, we use the term *imperative* as a classification of a sentence form type. The two most prominent morphosyntactic features of Hungarian imperatives are reversed verbal prefix (preverb) – verb stem order, and subjunctive morphology. In harmony with Varga (2013), we assume that imperatives, due to their full paradigms, are in the subjunctive mood. The paper investigates five types of imperatives: in addition to the basic type, it analyzes sentences where subjunctive morphology is combined with lengthened intonation, the hortative marker *hadd* ‘let’, and the discourse particles *csak* ‘just’ and *már* ‘already’.

From the functional perspective, imperatives exhibit great variation, which is recorded by Kaufmann (2012: 14) as the *functional inhomogeneity problem*. Besides the prototypical command/request interpretation, several possible speaker attitudes can lie behind an imperative sentence, for instance, concession, advice, threatening, asking for or giving permission – some of them are rather far from the meaning of the baseline imperative. Since our aim is to provide pragmatico-semantic analysis, our main interest

---

\* kleiber.judit@pte.hu

\*\* alberti.gabor@pte.hu

\*\*\* szabo.veronika@pte.hu

1 The present scientific contribution is dedicated to the 650th anniversary of the foundation of the University of Pécs, Hungary.

lies in the information states of the interlocutors, primarily the internal world of the speaker (addresser): his/her beliefs, desires and intentions. In the formal interpretation system we apply, it is possible to formalize this information, and then evaluate – through pattern matching between linguistic forms and world models – not only the propositional content of the sentence but also such pragmatic factors as speaker attitude or the sincerity of an utterance.

The paper is organized as follows. Section 2 presents the data: the above mentioned five Hungarian imperative sentence types. The theoretical framework, ReALIS, is introduced briefly in Section 3. Section 4 is devoted to the analysis of the data: the description of the intensional profiles of the baseline imperative, and then the four additional imperative sentence types, which can be regarded as the variants of the basic type fine-tuned by extra elements. Finally, Section 5 draws some conclusions.

## 2. THE DATA: FIVE HUNGARIAN IMPERATIVE SENTENCE TYPES

In this section, we present the data for five types of Hungarian imperatives, starting with the baseline followed by four types modified by extra elements. In Hungarian, an imperative sentence can appear in any person and number; this is the reason why our main (a) examples are all in the general, 3<sup>rd</sup> person singular form. With almost every type, however, we can point out the most common form, which indicates the preferred agent for the given imperative act (b examples), as well as the least common one with a (strongly) dispreferred agent, (c).

The formal features of the basic type include (i) the shared characteristics of all imperatives: subjunctive morphology and reversed preverb–verb order (if there is a preverb); as well as (ii) standard falling intonation typical for imperatives; and finally (iii) the lack of discourse markers. In the case of the baseline imperative (1), the preferred agent is the addressee (1b), and the dispreferred agent is the addresser (1c). Sentences like (1c) can only appear under special circumstances, for instance, when someone is talking to himself/herself.

- (1) a) Költözzön Péter Marihoz!  
 move.Sbjv.3Sg Péter Mari.Ade  
 ‘Péter should move to Mari’s.’  
 b) Költözz Marihoz!  
 move.Sbjv.2Sg Mari.Ade  
 ‘Move to Mari’s!’  
 c) \*Költözzek Marihoz!  
 move.Sbjv.1Sg Mari.Ade

As for their function, even the baseline imperative can express many speaker attitudes. Nevertheless, we can describe the prototypical use of the basic imperative as follows. On the one hand, the addresser (AR) longs for something, typically a change

in the present state of affairs. On the other hand, he/she expects cooperative behaviour from the addressee (ae). In the preferred 2sg case (1b), it means that ae should execute the action himself/herself (ae should move to Mari's); while in (1a), AR merely asks ae to have the same intentions as are his/hers (that Péter should move to Mari's). The general 3sg case constitutes a non-addressee-oriented directive (Péteri 2012) where ae plays a mediating, supporting or tolerating role.

Let us see the first variation of the basic type: an imperative sentence where the first syllable of the verb is unusually lengthened (2).

- (2) a) *Köööltözzön* Péter Marihoz!  
 move.Sbjv.3Sg Péter Mari.Ade  
 'Fine, Péter can move to Mari's, I do not mind (anymore).'  
 b) *Köööltözz* Marihoz!  
 move.Sbjv.2Sg Mari.Ade  
 'Fine, you can move to Mari's, I do not mind (anymore).'  
 c) *\*Köööltözzek* Marihoz!  
 move.Sbjv.1Sg Mari.Ade

The effect is that now it is not AR who longs for the given action (he/she is rather against its coming true), but someone else: typically ae, or perhaps the Agent of the action (Ag). AR's intentions are somewhat neutral, he/she merely consents to it. The given speech act, thus, expresses *concession*. Similarly to the baseline imperative, the preferred agent is ae (2b), while it is excluded that AR, who gives the permission, and Ag, who obtains it, coincide (2c).

The third type contains the hortative marker *hadd* 'let' (3). Contrary to the previous types, the preferred agent is AR (3b), and the dispreferred agent is ae (3c).

- (3) a) *Hadd* költözzön Péter Marihoz!  
 let move.Sbjv.3Sg Péter Mari.Ade  
 'Let Péter move to Mari's.'  
 b) *Hadd* költözzek Marihoz!  
 let move.Sbjv.1Sg Mari.Ade  
 'Let me move to Mari's.'  
 c) *\*Hadd* költözz Marihoz!  
 let move.Sbjv.2Sg Mari.Ade

In harmony with Szücs (2010) and Gärtner–Gyuris (2012), we can establish that, in sentences with subjunctive verb morphology, two main meaning components can be assigned to the Hungarian hortative marker *hadd* 'let'. The first one can be paraphrased as 'ask for permission', and it can appear with 1st and 3rd person action verb forms (3a–b). The second one means 'grant permission', and it only occurs with 3rd-person verbs (3a). In this paper, we only discuss the former usage. With this type, it is ae who is assumed not to long for the move, unlike AR and/or the Agent. In the preferred 1sg

case (3b), AR conveys his/her own desires; while in the general 3sg case (3a), AR could merely be an “advocate” for Ag (Péter).

The next type of imperative contains the discourse particle *csak* ‘just’ (4). As with most types, the preferred agent is ae (4b), and the dispreferred one is AR (4c). This time, however, sentences like (4c) do exist, they merely have a slightly different meaning.

- (4) a) Költözzön      *csak* Péter Marihoz!  
       move.Sbjv.3Sg just Péter Mari.Ade  
       ‘Let Péter move to Mari’s. (Let him try and do that.)’
- b) Költözz      *csak* Marihoz!  
       move.Sbjv.2Sg just Mari.Ade  
       ‘You should just move to Mari’s. (Just try and do that.)’
- c) Költözzek    *csak* Marihoz!  
       move.Sbjv.1Sg just Mari.Ade  
       ‘Just wait until I move to Mari’s! (You will see what happens.)’

Combined with subjunctive morphology, the particle *csak* ‘just’ can express several speaker attitudes: threatening, hastening, encouraging (Fábricz 1986: 78), depending on such contextual factors as intonation and the agent of the action, among others. This paper only discusses the most prominent, “sinister” usage of *csak* ‘just’. With this type, AR is definitely against the move, which is now assumed to be longed for very much by Ag (or, perhaps, by ae). On the surface, AR’s intention is neutral, he/she does not want (for ae) to prevent the action. With 2nd-person verb forms (4b), this can be paraphrased as: ‘I am tired of persuading you, just do what you want, but you will see ...’. It is very likely, though, that the speaker’s intention is, in fact, the exact opposite, and the utterance is a final effort to persuade the interlocutor *not* to do it (“reversed psychology”). The dispreferred (4c) form also reflects this sinister attitude: AR predicts that this change would be unpleasant for ae (or someone else).

The last discussed type contains the particle *már* ‘already’ (5). The preferred agent is ae (5b); while the dispreferred one is AR (5c), occurring only when AR is talking to him/herself.

- (5) a) Költözzön      *már* Péter Marihoz!  
       move.Sbjv.3Sg already Péter Mari.Ade  
       ‘I want Péter to decide to move to Mari’s at long last.’
- b) Költözz      *már* Marihoz!  
       move.Sbjv.2Sg already Mari.Ade  
       ‘Move to Mari’s already!’
- c) \*Költözzek    *már* Marihoz!  
       move.Sbjv.1Sg already Mari.Ade

The particle *már* ‘already’ can be associated with a wide range of speaker attitudes in imperative sentences, such as hastening, threatening, encouraging, persuading or

begging (Fábricz 1986: 188). In its most common usage, it expresses *hastening*, and the fact that the change would be desirable (Fábricz 1986: 70). In (5a–b), AR thinks that someone, preferably Ag, longs for the move very much, and (hence) wants this person to realize his/her wishes.

This section presented the forms and the primary functions of five Hungarian imperative sentence types and elaborated on their pragmatic-semantic interpretations. The examples have demonstrated that imperatives can be associated with various kinds of distribution of desires and intentions among the three interested participants: the addresser, the addressee and the Agent of the action. In the following two sections, we introduce the framework in which we then provide formal analysis for the discussed imperative sentence types. Our aim is to represent the interlocutors' beliefs, desires and intentions so that the changes in speaker attitudes could be captured by different parameter settings.

### 3. THE FRAMEWORK: SOME WORDS ON $\Re$ ALIS

This section briefly describes the applied theoretical framework  $\Re$ ALIS 'Reciprocal And Lifelong Interpretation System' (Alberti 2011). It would go beyond the scope of this short paper to enumerate arguments for  $\Re$ ALIS (the interested reader is referred to Alberti and Kleiber (2012), Alberti (2012), and Alberti and Nóthig (2015)); thus, we only highlight one distinctive property and then provide a short introduction to the formalism.

$\Re$ ALIS can be characterized as a discourse-representation-based (Kamp et al. 2011; Asher and Lascarides 2003) formal semantic theory with a radically new ontology (Alberti and Kleiber 2014). Our starting point is that, in order to account for pragmatic phenomena, we should be able to examine not only the outside world but also the interlocutors' internal worlds (mental states): their beliefs, desires and intentions. The innovative feature of  $\Re$ ALIS is that (all) representations are regarded as mental states, which are taken to be part of the world model. Thus, this approach eliminates the "extra level" between the world model and the linguistic form, which is considered to be problematic by the Amsterdam School (e.g., Groenendijk and Stokhof 1991).

The motivation behind founding  $\Re$ ALIS on the basis of this ontology is to create a homogeneous structure for the three representational levels needed to account for linguistic phenomena: the representation of (1) the discourse, (2) the world, and (3) the human mind. In this way, we can decide the truth value of a proposition which exists solely in a person's mind with the same pattern-matching mechanism applied for evaluating utterances referring to the outside world. Due to the fact that all three types of information are represented in the same structure, their formal examination and comparison is possible. For instance, a mismatch between the outside world and the discourse suggests some kind of mistake, while a mismatch between the discourse and the speaker's mind indicates a form of deceit: lie, bluff, white lie, and so on, depending on the parameters of the deviation (Alberti, Vadász and Kleiber 2014).

In  $\Re$ ALIS, it is crucial to differentiate between the addresser/addressee and the speaker/listener roles. The former belongs to the ideal case (cf. Searle 1969), the

linguistically encoded information of an utterance, while the latter appears in a concrete situation which may not realize the ideal case. For instance, a promise may be insincere, the speaker could be dishonest, the listener may not recognize the irony, and so on. During the interpretation process, it is to be evaluated from clause to clause – in harmony with Oishi’s thesis (2014) – whether the speaker is acting legitimately, sincerely, and/or adequately, while, in the on-going discourse, playing the addresser’s role and giving the listener the addressee’s role. With this approach, when a proposition is evaluated against the current content of the interlocutors’ information state, various pragmatic factors can be accounted for, such as the Gricean maxims – e.g., the sincerity or the relevance of an utterance –, irony, politeness, and so on.

In the remainder of the section, we briefly introduce the applied formalism. In *ReALIS*, a clause performed in an on-going discourse conveys a piece of information which belongs to an *intensional profile*. For instance, the sentence in (1a) is assigned an intensional profile which characterizes the baseline imperative by representing its pragmatic-semantic contribution: the interlocutors’ beliefs, desires and intentions when performing the utterance.

An intensional profile consists of finite components of *worldlets*. A worldlet encodes one meaning component, such as a desire for an eventuality, or a belief about the intentions of our partner. It can be regarded as a labeled DRS where a *level label* encodes five essential properties concerning the given piece of information: 1. modality (M): Belief, Desire or Intention; 2. intensity (I) of the modality: some, great, (almost) Maximal, allowing multiple values; 3. host (R) of the worldlet: primarily AR or ae; 4. time parameter (T); and 5. polarity value (P): + (true), – (false), 0 (not specified), also allowing multiple values. The label  $\langle B, M, AR, \tau, + \rangle$ , for instance, represents that the addresser (AR) knows (Maximally Believes) at time  $\tau$  that a given eventuality  $e$  holds (+). A worldlet can be embedded in another worldlet which makes it possible to refer to information states. For instance, the series of level labels  $\langle B, M, AR, \tau, + \rangle \langle D, M, ae, \tau, + \rangle$  assigned to a worldlet encodes that AR is sure that  $ae$  long for  $e$ .

The last discussed feature of *ReALIS* is that a piece of information can appear in several worldlets simultaneously, which can be understood as a *prism effect*. When eventuality  $e$  is represented in the interlocutor’s mind, it is “scattered” like a prism scatters the light (hence the term). The analysis will demonstrate this effect shortly. Thus, an intensional profile is an element of the set  $P((M \times P(I) \times R \times T \times P(P))^*)$ : the power set of the set of finite sequences of level labels. A clause is interpreted against a worldlet in order to obtain its truth conditions and other semantic and/or pragmatic well-formedness conditions in the given context.

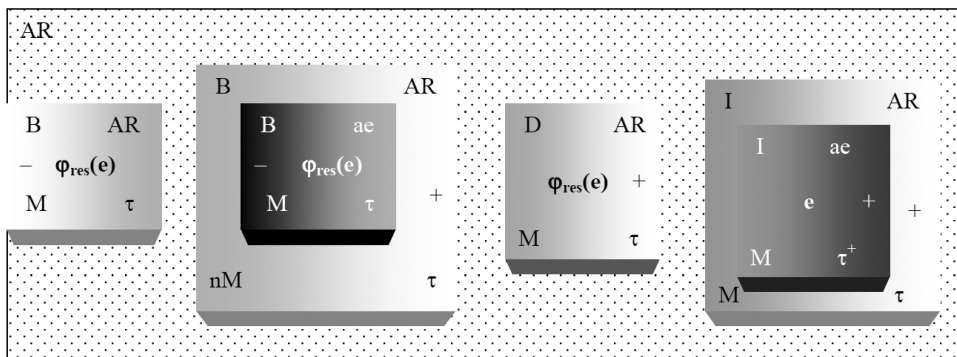
#### 4. THE ANALYSIS: THE INTENSIONAL PROFILES OF THE DISCUSSED IMPERATIVES

In this section, we present the intensional profiles of the five types of Hungarian imperatives discussed in Section 2 starting with the baseline imperative ((1a), repeated as (6a)). The eventuality  $e$  expressed by (6a), namely Péter’s moving to Mari’s, appears

in four different worldlets (prism effect) which are displayed in (6b–e) along with their paraphrases. It is followed by the visual representation of the profile in Figure 1 which is essentially a conglomerate of (S)DRS boxes.

- (6) a) Költözzön Péter Marihoz!  
 move.Sbjv.3Sg Péter Mari.Ade  
 ‘Péter should move to Mari’s.’
- b)  $\langle B, M, AR, \tau, - \rangle$  “I (AR) am sure that the result phase  $\varphi_{res}(e)$  of the given eventuality  $e$  does not hold (polarity:  $-$ ) (i.e., Péter and Mari still live in different flats, that is, Péter has not moved to Mari’s yet)”
- c)  $\langle B, nM, AR, \tau, + \rangle \langle B, M, ae, \tau, - \rangle$  “I think that you ( $ae$ ) are also aware of this fact (the certainty of the assumption is ‘ $nM$ ’=‘non-maximal’).”
- d)  $\langle D, M, AR, \tau, + \rangle$  “I long for the aforementioned result phase.”
- e)  $\langle I, M, AR, \tau, + \rangle \langle I, M, ae, \tau^+, + \rangle$  “I want you to intend the action, too (at a later point  $\tau^+$  in time).”

Figure 1: The intensional profile of the baseline Hungarian imperative sentence type



So the intensional profile of the baseline imperative constitutes of four worldlets. The first two worldlets are concerned with AR’s beliefs (6b–c); the third worldlet represents AR’s desire (6d); and the last worldlet encodes AR’s intention (6e). In the prototypical 2sg case ((1b) in Section 2), acting in favour of this intention should lead to  $ae$ ’s executing  $e$  him/herself; while in the general 3sg case (6a),  $ae$  is assigned a mediating or supporting role.

Let us now turn to the intensional profiles of the four additional imperative sentence types we discussed in Section 2. The examination reveals that there are rather few parameters which differ from the basic setup – in contrast to the major differences between, for instance, the basic declarative and the basic imperative profiles (Alberti and Kleiber 2014). Thus, we can refer to the four additional types as the variants of the basic type fine-tuned by extra elements. In what follows, we elaborate on the parametric changes responsible for different speaker attitudes associated to imperative types.



It is common in all types of imperatives that AR is sure that  $\varphi_{\text{res}}(e)$  does not hold (Péter has not moved to Mari's yet) and assumes that ae is also aware of this fact. If there is no such shared piece of knowledge, the speaker has illegitimately given the listener the addressee role.<sup>2</sup> So there is no variation among the different types regarding AR's beliefs.

As for intentions, we have illustrated in Section 2 that basically two types of intention pattern are relevant to imperatives: an active one, when AR actually wants e to happen, and a neutral/passive one, when he/she does not oppose it. The worldlet which encodes the first shade of meaning  $\langle I, M, AR, \tau, + \rangle \langle I, M, ae, \tau^+, + \rangle$  appears in the intensional profile of the basic type (6e) and the one with the particle *már* 'already' expressing hastening (example (5) in Section 2). Changing the polarity values to negative  $\langle I, M, AR, \tau, - \rangle \langle I, M, ae, \tau^+, - \rangle$  results in the second shade of meaning: AR asks ae not to prevent e (negative intention = prevention; AR's intention is to prevent ae's prevention). This attitude is represented in the other three cases.

Table 1 below provides a comparative overview of the discussed types. The first two rows represent AR's beliefs concerning  $\varphi_{\text{res}}(e)$  or ae's knowledge. The third row encodes the relevant desires (to be elaborated on). The fourth row displays AR's intentions we have presently discussed. The last three rows encode additional information about the uses and interpretations of these sentences: preferred and dispreferred agents (discussed in Section 2) as well as the preferred identification for  $r^*$  occasionally occurring in the descriptions.

Table 1: A comparison between the intensional profiles of the five Hungarian imperative sentence types discussed in the paper

	a. Basic	b. CVVVC...	c. <i>hadd</i>	d. <i>csak</i>	e. <i>már</i>
AR's knowledge conc. $\varphi_{\text{res}}(e)$	$\langle B, M, AR, \tau, - \rangle$	←	←	←	←
ae's knowledge conc. $\varphi_{\text{res}}(e)$ (acc. to AR)	$\langle B, nM, AR, \tau, + \rangle$ $\langle B, M, ae, \tau, - \rangle$	←	←	←	←
AR's, ae's and/or Ag's desire conc. $\varphi_{\text{res}}(e)$	$\langle D, M, AR, \tau, + \rangle$	$\langle D, M, AR, \tau, 0- \rangle$	$\langle D, M, AR, \tau, 0+ \rangle$	$\langle D, nM, AR, \tau, - \rangle$	$\langle D, M, AR, \tau, 0+ \rangle$
		$\langle B, nM, AR, \tau, + \rangle$ $\langle D, M, r^*, \tau, + \rangle$	$\langle B, nM, AR, \tau, + \rangle$ $\langle D, M, ae, \tau, 0- \rangle$	$\langle B, nM, AR, \tau, + \rangle$ $\langle D, M, r^*, \tau, + \rangle$	$\langle B, nM, AR, \tau, + \rangle$ $\langle D, M, r^*, \tau, + \rangle$
			$\langle B, nM, AR, \tau, + \rangle$ $\langle D, M, r^*, \tau, + \rangle$		$\langle D, M, AR, \tau, + \rangle$ $\langle I, M, r^*, \tau, + \rangle$
AR's intention conc. e and/or ae's intention	$\langle I, M, AR, \tau, + \rangle$ $\langle I, M, ae, \tau^+, + \rangle$	$\langle I, sm, AR, \tau, 0 \rangle$	$\langle I, M, AR, \tau, - \rangle$ $\langle I, M, ae, \tau^+, - \rangle$	$\langle I, M, AR, \tau, 0- \rangle$ $\langle I, M, ae, \tau^+, - \rangle$	$\langle I, M, AR, \tau, + \rangle$ $\langle I, M, ae, \tau^+, + \rangle$ $\langle I, M, r^*, \tau^+, + \rangle$
		$\langle I, M, AR, \tau, - \rangle$ $\langle I, M, ae, \tau^+, - \rangle$			
Pref'd.: Ag=	ae	ae	AR	ae	ae
Dispr'd.: Ag≠	AR	AR	ae	AR	AR
Pref'd.: $r^*=$		ae > Ag	AR > Ag	Ag > ae	Ag > ae

2 This is an excellent example of cases when evaluation against a model does not serve the purpose of obtaining a truth value on the basis of the model of the real world but serves the purpose of calculating pragmatic felicity conditions on the basis of the content of the interlocutors' minds.

Among the imperatives we have examined, desires have exhibited the greatest variation (see the third row in Table 1). Except for the baseline where only AR's – positive – desire is encoded toward  $\varphi_{\text{res}}(\text{e})$ , (6d), the imperative profiles contain AR's own desires as well as AR's assumptions concerning ae's – or occasionally someone else's (e.g., Ag's) – desires.

First, let us examine the intensional profile of the imperative with the unusually lengthened first syllable (CVVVC..., column b in Table 1) introduced in (2) in Section 2 which was assigned the speech act of *concession*. It contains two worldlets concerning desires (third row of Table 1):  $\langle \text{D}, \text{M}, \text{AR}, \tau, 0- \rangle$  expressing that AR does not long for  $\varphi_{\text{res}}(\text{e})$  (0-: neutral or opposing stance); and  $\langle \text{B}, \text{nM}, \text{AR}, \tau, + \rangle \langle \text{D}, \text{M}, \text{r}^*, \tau, + \rangle$  expressing that AR assumes that a person r\* longs for it, where r\* preferably coincides with ae or Ag, in this order (last row).

The next type we discussed contained the hortative marker *hadd* 'let' (3), and it was assigned the speech act of *asking for permission*. Its intensional profile (column c in Table 1) encodes that according to AR ae does not long for  $\varphi_{\text{res}}(\text{e})$  ( $\langle \text{B}, \text{nM}, \text{AR}, \tau, + \rangle \langle \text{D}, \text{M}, \text{ae}, \tau, - \rangle$ ), while AR and Ag, who preferably coincide (fifth row), do. The 0 in  $\langle \text{D}, \text{M}, \text{AR}, \tau, 0+ \rangle$  refers to the case when r\*=Ag (last row), that is when AR conveys someone else's wishes.

The intensional profile of the "sinister" usage of *csak* 'just' is displayed in column d in Table 1, expressing some kind of *threatening*. The polarity value of AR's desire is negative ( $\langle \text{D}, \text{nM}, \text{AR}, \tau, - \rangle$ ), while r\* (= Ag or ae) is assumed to have strong positive desire for  $\varphi_{\text{res}}(\text{e})$ .

Finally, in the last (e) column of Table 1 the intensional profile for the speech act of *hastening* is presented which is assigned to the imperative type containing the discourse marker *már* 'already'. It encodes that AR has either positive or neutral desire for  $\varphi_{\text{res}}(\text{e})$  ( $\langle \text{D}, \text{M}, \text{AR}, \tau, 0+ \rangle$ ) while believing that r\* (preferably=Ag=ae) really longs for it. AR also wishes to get r\* (ae) to intend to do e ( $\langle \text{D}, \text{M}, \text{AR}, \tau, + \rangle \langle \text{I}, \text{M}, \text{ae}, \tau, + \rangle$ ) via helping this person realize his/her wishes. The three-component worldlet in the "Intentions" row becomes the usual two-component one when applied to r\*=Ag=ae; while in the case of r\*=Ag $\neq$ ae (3rd-person verb forms) it encodes the function when ae plays a mediating role.

In this section, we have presented the intensional profiles of the five Hungarian imperative sentence types introduced in Section 2. The analysis has demonstrated that changing a few parameters – polarity values for the interlocutors' intentions and/or desires – can account for the different speech acts associated with different types of imperatives.

A possible direction for future research would address the notion of compositionality in order to answer questions like how the intensional profiles of imperatives are derived and what role the discourse markers play in the process. For the time being, ReALIS applies two means for providing compositional analysis: the simple operation of concatenation, on the one hand, which is suitable for deriving the pragmatic-semantic contribution (intensional profiles) of interrogative imperatives (a basic imperative modified by interrogative elements); and the formal operation of semantic blending

(based on the cognitive linguistic notion), on the other hand, which is capable of mixing partially incompatible meaning components, such as mood and modality (Alberti, Dóla and Kleiber 2014). Furthermore, we need to provide underspecified representations for discourse markers which could then be composed with the baseline imperative profile in order to acquire a compositional analysis.

## 5. CONCLUSION

The paper investigated the pragmatic-semantic contributions of various imperative sentence types in Hungarian. We have examined the baseline imperative which is prototypically used for commanding, along with four additional types modified by extra elements: imperative sentences with an unusually lengthened syllable expressing concession, the hortative marker *hadd* ‘let’ expressing permission, the discourse marker *csak* ‘just’ expressing some kind of threatening, and the discourse marker *már* ‘already’ expressing hastening. We have analyzed these types in the discourse-semantic framework  $\Re$ ALIS and provided their intensional profiles which represent the interlocutors’ beliefs, desires and intentions in a formal manner.

We have established that basically two types of intention pattern and several distributions of desires – between the addresser, the addressee and the Agent of the action – play a part in capturing the pragmatic-semantic contributions of different types of imperatives associated with different speech acts. The analysis has demonstrated that the different meanings of the variants can be derived from parametric differences.

The advantageous innovation of  $\Re$ ALIS can be formulated as follows: checking whether the speaker and the hearer are suitable for serving as the addresser and the addressee of the linguistically defined speech act simply requires a truth-conditional investigation primarily into certain worldlets in the addresser’s mind (e.g., what (s)he hypothesizes and longs for, and what (s)he assumes that certain other persons hypothesize and long for). The task boils down to get to the worldlets in which certain polarity values must then be checked.

## References

- ALBERTI, Gábor (2011)  *$\Re$ ALIS: Interpretálók a világban, világok az interpretálóban* [Interpreters in the world, worlds in the interpreter]. Budapest: Akadémiai Kiadó.
- ALBERTI, Gábor (2012) “Where are Possible Worlds? II. Pegs, DRs, Worldlets and Reification.” In: Alberti G. et al. (eds), *Vonzásban és változásban*. Pécs: Doctoral School of Linguistics at University of Pécs, 308–323.
- ALBERTI, Gábor/Mónika DÓLA/Judit KLEIBER (2014) “Mood and modality in Hungarian: Discourse Representation Theory meets Cognitive Linguistics.” *Argumentum* 10, 172–191.
- ALBERTI, Gábor/Judit KLEIBER (2012) “Where are Possible Worlds? Arguments for  $\Re$ ALIS.” *Acta Linguistica Hungarica* 59, 3–26.

- ALBERTI, Gábor/Judit KLEIBER (2014) “ReALIS: Discourse Representation with a Radically New Ontology.” In: L. Veselovská/M. Janebová (eds), *Complex Visibles Out There*. Olomouc Modern Language Series 4. Olomouc: Palacký University, 513–528.
- ALBERTI, Gábor/László NÓTHIG (2015) “ReALIS2.1: The Implementation of Generalized Intensional Truth Evaluation and Expositive Speech Acts in On-Going Discourse.” *International Journal on Advances in Intelligent Systems* 8/1&2, 85–106.
- ALBERTI, Gábor/Noémi VADÁSZ/Judit KLEIBER (2014) “Ideal and Deviant Interlocutors in a Formal Interpretation System.” In: A. Zuczkowski et al. (eds), *The communication of certainty and uncertainty*. Amsterdam: Benjamins, 59–78.
- ASHER, Nicholas/Alex LASCARIDES (2003) *Logics of Conversation*. Cambridge: Cambridge Univ. Press.
- FÁBRICZ, Károly (1986) *Partikulák a magyar és az orosz nyelvben* [Particles in Hungarian and Russian]. Kandidátusi értekezés. Szeged.
- GÄRTNER, Hans-Martin/Beáta GYURIS (2012) “Pragmatic markers in Hungarian: Some introductory remarks.” *Acta Linguistica Hungarica* 59, 387–426.
- GROENENDIJK, Jeroen/Martin STOKHOF (1991) “Dynamic Predicate Logic.” *Linguistics and Philosophy* 14, 39–100.
- KAMP, Hans/Joseph van GENABITH/Uwe REYLE (2011) “Discourse Representation Theory.” In: D. Gabbay/F. Guenther (eds), *Handbook of Philosophical Logic*. Berlin: Springer, 125–394.
- KAUFMANN, Magdalena (2012) *Interpreting Imperatives*. Dordrecht: Springer.
- OISHI, Etsuko (2014) “Discursive functions of evidentials and epistemic modals.” In: S. Cantarini et al. (eds), *Certainty-uncertainty – and the attitudinal space in between*. Amsterdam: Benjamins, 239–262.
- PÉTERI, Attila (2012) “The Hungarian Imperative Particle *Hadd*.” *Acta Linguistica Hungarica* 59, 439–463.
- SEARLE, John (1969) *Speech Acts*. Cambridge: Cambridge University Press.
- SZÜCS, Márta (2010) “A *hadd* problémaköre [The problem of *hadd* ‘let’].” *LingDok* 9, 193–210.
- VARGA, Diána (2013) *A magyar felszólító mondatok szerkezete* [The structure of Hungarian imperative sentences]. PhD dissertation. Pázmány Péter University.

Summary  
THE INTENSIONAL PROFILES OF FIVE HUNGARIAN  
IMPERATIVE SENTENCE TYPES

The paper investigates Hungarian imperative sentence types from a pragmatic-semantic point of view. In addition to the baseline imperative, it analyzes types with non-standard intonation pattern and/or discourse markers. We apply the (S)DRT-based discourse-semantic framework of ReALIS for the examination and representation of

five form–function pairs. The discussed types are all assigned an *intensional profile* which encodes the interlocutors' beliefs, desires and intentions (from the addresser's perspective). The analysis derives the different meanings of the variants from parametric differences.

**Keywords:** discourse marker, discourse representation structure, imperative sentence type, intensionality

Povzetek  
INTENZIONALNI PROFILI PETIH MADŽARSKIH  
VELELNIH STAVČNIH TIPOV

Članek obravnava madžarske velelne stavke z vidika pragmatike in semantike. Poleg nezaznamovanega velelnika analiziramo tudi tipe velelnikov z nestandardnimi intonacijskimi vzorci in/ali diskurzniimi označevalci. Z diskurzno-semantičnim okvirom ReALIS, ki temelji na (S)DRT, raziščemo in predstavimo pet oblikovno-funkcijskih parov. Obravnavanim tipom določimo *intenzionalni profil*, ki obsega prepričanja, želje in namere sogovornikov (z vidika govorca). Analiza na podlagi parametričnih razlik izpelje različne pomene obravnavanih variant.

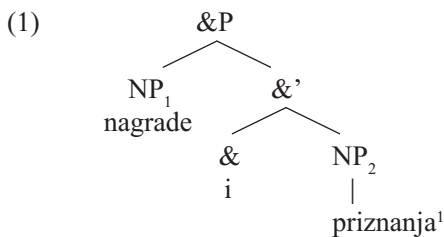
**Ključne besede:** diskurzni označevalec, struktura reprezentacije diskurza, velelni stavki, intenzionalnost



## STRATEGIES OF PARTICIPLE AGREEMENT WITH CONJOINED SUBJECTS IN BOSNIAN/CROATIAN/SERBIAN

### 1. INTRODUCTION

In Bosnian/Croatian/Serbian (BCS), the agreement forms of the participle in the predicate may be realised in three different ways when the subject consists of conjoined noun phrases depending on the phi features of these noun phrases. The participle may agree in number and gender with the conjoined subject phrase as a whole (agreement with the maximal projection, i.e. projection of the conjunction (&) as the head of a conjoined phrase (&P)), or it may agree with the hierarchically highest noun phrase (NP<sub>1</sub>), or it may agree with the noun phrase closest to the participle (NP<sub>2</sub>), as represented in (1) for the conjoined phrase *nagrade i priznanja* ‘awards and recognitions’:



The experimental research we present in this paper demonstrates that the participle may be realised in three different forms, as in the following examples, which illustrate participle agreement forms when the subject consists of conjoined plural noun phrases of different genders [Feminine + Neuter]:

\* nedzad.leko@ff.unsa.ba

1 We assume that conjoined NPs have a hierarchical structure, as presented in (1). However, it should be pointed out that other opinions can be found in the literature – there are assumptions that the configuration of these phrases is linear (flat) (see Culicover and Jackendoff 2005: 143) even in narrow syntax.

(2) a) *Nagrade i priznanja su uručeni studentima.*  
 awards<sub>F</sub> and recognitions<sub>N</sub> AUX handed<sub>M</sub> students  
 ‘The awards and recognitions were handed to the students.’

b) *Nagrade i priznanja su uručena studentima.*  
 handed<sub>N</sub>

c) *Nagrade i priznanja su uručene studentima.*  
 handed<sub>F</sub>

The masculine form (*uručeni*) in (2a) illustrates agreement with the maximal projection (&P) in (1), which is a projection of the conjunction & (*i*) as the head of a conjoined phrase. The conjoined phrase (&P) is masculine in gender by default, and consequently, the participle in (2a) does not agree either with NP<sub>1</sub> or NP<sub>2</sub>, but rather with the maximal projection (&P). The neuter form (*uručena*) in (2b) illustrates agreement with NP<sub>2</sub> (*priznanja*), which is closest to the participle. Finally, the feminine form in (2c) illustrates agreement with NP<sub>1</sub> (*nagrade*), which is hierarchically in the highest position in this conjoined phrase.

These data point to the existence of three distinct strategies of subject-predicate agreement when the subject consists of conjoined noun phrases. Gender and number agreement forms of participles in the predicate may be computed in three different ways, depending on the features of conjoined noun phrases in the subject. Participles may agree in gender and number with the subject phrase as a whole (that is agreement with the maximal projection – a Boolean Phrase (&P)), or with the conjunct which is closest to the participle, or with the conjunct which is hierarchically the highest conjunct. We demonstrate in this paper that, in this respect, BCS behaves like Slovenian, which also has three strategies of agreement, as outlined by Marušič et al. (2015).

In order to test the initial hypothesis that there are three agreement strategies, we conducted a controlled experimental study of the morphosyntax of agreement between conjoined subjects and participles in BCS involving oral and written production experiments. The first experiment was a computer-assisted oral experiment in which the participants in the experiment were given first a test sentence like *Poklon je uručen na pozornici* ‘The present was handed at the stage’, and then conjoined noun phrases like *nagrade i priznanja* ‘awards and recognitions’. Their task was to insert this phrase in the sentence instead of the noun *poklon* ‘present’, which would involve changing the participle agreement form. Both word orders were tested – the subject preceding the predicate (SV order), as in the previous example, and the opposite order (VS), as in *Na pozornici je uručen poklon*. The second experiment was a written experiment, which was conducted after the oral one. The written test contained exactly the same material as the oral experiment, with the same distribution of sentences, and the same participants were involved in both the oral and the written experiments.

The experiments revealed a high presence of default, masculine agreement and closest conjunct agreement. 50% of preverbal conjoined phrases elicited the default

masculine agreement, and 95% of postverbal conjoined phrases elicited the closest conjunct agreement. However, the relatively high percentage of participle forms agreeing with the first conjunct of the preverbal conjoined phrase confirmed that highest conjunct agreement (HCA) is a legitimate agreement strategy in BCS. On the other hand, a very small number of participle forms agreeing with the last conjunct of the postverbal conjoined phrase confirmed that last conjunct agreement (LCA) in postverbal contexts cannot be recognised as a separate agreement strategy. Instead, such examples should be characterised as performance errors. These results are contrary to Bošković's findings (2009), in which he does not acknowledge HCA as a legitimate strategy, however, our results do confirm the findings of Marušič et al. (2015).

The paper is organised as follows. In section 2, we briefly discuss previous accounts of predicate agreement with conjoined subject noun phrases; in section 3, we present the experimental methodology; in section 4, we discuss agreement with uniform gender conjuncts; in section 5, we provide a detailed analysis of agreement with mixed gender conjuncts; in section 6, we discuss our findings, summarising the main results.

## **2. PREVIOUS ACCOUNTS**

The type of predicate agreement with conjoined subject noun phrases has been intensively studied both theoretically and experimentally, and, particularly in Slavic languages, it has received special attention. These studies can certainly contribute to a wider theoretical discussion about the role of agreement in grammar (this is discussed, for example, by Benmamoun et al. (2010), Munn (1999), Bhatt and Walkow (2013)), as well as to experimental investigations about attraction phenomena (as discussed, for example, by Bock and Miller (1991), Franck et al. (2006, 2007), and Franck (2011)). The investigations of agreement phenomena in Slavic languages are dominated by two approaches. One approach is exclusively syntactic, in which the syntactic analysis of agreement phenomena is based on native speaker intuitions or theoretical predictions (this kind of approach is taken by Bošković (2009) and Puškar and Murphy (2015)). The second approach, on the other hand, may be characterised as multi-component, arguing that in addition to the syntactic component, agreement processing involves also another component (phonetic) at the level of Phonetic Form (PF); this kind of approach is taken by Marušič et al. (2007, 2015).

### **2.1. Bošković (2009)**

Bošković (2009) offers a uniform account of first and last conjunct agreement based on the operation Agree. According to Bošković, with postverbal subjects, participles in Serbo-Croatian always exhibit first conjunct agreement (for gender), and with preverbal subjects, only last conjunct agreement (for gender) is exhibited. His basic assumption is that the agreement is handled exclusively in the syntax by the operation Agree. In his analysis, Bošković proposes that the probe that is responsible for participial agreement searches for a goal to value its number and gender features. Since Conjunction Phrase (&P) is specified only for number, the probe finds disjoint valuators, &P for number



and the first conjunct for gender. This happens in cases of first conjunct agreement (with postverbal subjects). Bošković claims that the probe is a single probe (as opposed to approaches which argue for separate probes).

However, in the (abstract) structure which leads to last conjunct agreement, the Part (participle) probes for phi features, matching &P for number and NP<sub>1</sub> (the first conjunct) for gender. Since Part has an EPP feature, a phrase must move to SpecPartP. However, the problem is that there are two valuators, one requiring pied-piping of &P and the other requiring pied-piping of NP<sub>1</sub>. Since both &P and NP<sub>1</sub> are in principle pied-pipable in Serbo-Croatian (this language allows Left Branch Extraction (LBE) condition violation), Bošković suggests that this kind of ambiguity prevents pied-piping, and consequently the valuation itself is blocked. The participial probe then initiates a second probing operation within a larger search space that includes NP<sub>2</sub> (the second conjunct). The second conjunct can value the gender feature of the probe, and since it is in principle immobile, it is not a candidate for movement. Consequently, a valuator that undergoes pied-piping can be unambiguously determined. The Agree operation is then followed by movement of &P to SpecPartP. This happens in the case of last conjunct agreement with preverbal subjects.

## 2.2. Marušič et al. (2015)

In their paper *The Grammars of Conjunction Agreement in Slovenian* (2015), Marušič, Nevins, and Badecker document three agreement strategies in Slovenian after conducting five experimental studies. The agreement can target one of the three feature-bearing controllers: Conj<sub>1</sub>, Conj<sub>2</sub>, or BoolP. Consequently, the locality criterion does not only mean the choice between the hierarchically highest or the linearly closest conjunct. It can also mean agreement with the closest phrase (XP) of the relevant type (BoolP). The factor that affects the gender feature computation process is that masculine is an unmarked gender in Slovenian (as in BCS).

In one of the grammars of agreement in Slovenian that they describe, operation Agree targets the BoolP and does not probe at the individual conjuncts. Since in their view, a Conj head cannot compute its own gender value, what we have is the default insertion of the masculine value into phi features on the participle.

In another grammar of agreement, Agree targets the BoolP first. However, as it finds no gender value, rather than inserting the default masculine values, it continues to probe within the BoolP. Which conjunct is the source for the gender feature is a matter of locality: it is either the hierarchically highest or the linearly closest conjunct. The process of linearisation whereby the BoolP structure is flattened affects the choice here. If the Agree-Copy operation takes place before conjunct flattening, the hierarchically closest conjunct is the gender agreement controller. However, if Agree-Copy takes place after the BoolP structure has been flattened, the linearly closest conjunct is selected by the Probe.

Marušič et al. (2015) assume that operation Agree is carried out in two steps: Agree-Link and Agree-Copy. Agree-Link always applies in narrow syntax, but Agree-Copy can apply either in syntax or post-syntactically, and the authors additionally

assume that its order of application can vary with respect to linearisation (i.e. flattening of BoolP).

### 3. METHODOLOGY AND OVERVIEW OF THE EXPERIMENTS

As described in the Introduction, we demonstrate that BCS, like Slovenian, has three distinct strategies of subject-predicate agreement when the subject consists of conjoined noun phrases. Gender and number agreement forms of participles in the predicate may be computed in three different ways, depending on the features of conjoined noun phrases in the subject. Participles may agree in gender and number with the subject phrase as a whole (it is agreement with the maximal projection – a Boolean Phrase), or with the conjunct which is closest to the participle, or with the conjunct which is hierarchically the highest conjunct.

In order to prove this claim, we performed a controlled experimental study of the morphosyntax of agreement between conjoined subjects and participles in BCS. The experiments were conducted at the University of Sarajevo as a part of the EMSS project (*Experimental Morphosyntax of South Slavic Languages*, Leverhulme Trust/University College London). We conducted two types of studies: oral elicitation and written elicitation, and in both we registered variability in production elicited.

Both experiments – oral elicitation immediately followed by written elicitation – were administered as part of a single session to the same group of participants. All participants in the experiments were students at the University of Sarajevo who are native speakers of BCS from Sarajevo. There were 30 participants: all third-year students at the University of Sarajevo who had finished primary and secondary school in Sarajevo or the immediate region.

The first experiment was a computer-assisted oral experiment. Participants were recorded and prompted by a computer screen to continue to the next sentence. The experiment was administered individually using the online software Ibex (Drummond 2011). It involved a self-paced reading and sentence-completion task. Participants read a model sentence appearing on the screen. They then saw a replacement noun phrase and were asked to replace the subject of the model sentence with this new noun phrase. Responses were recorded using Audacity and coded afterwards according to their agreement features.

The experimental design involved six examples for each of the following nine gender combinations of plural conjuncts: [Masc + Masc], [Fem + Fem], [Neut + Neut], [Masc + Fem], [Masc + Neut], [Fem + Masc], [Fem + Neut], [Neut + Masc], and [Neut + Fem]. Additionally, the experimental material involved fillers. There were three filler conditions: paucal (numerals 2, 3, 4) with a head noun in masculine singular, a hybrid noun<sup>2</sup> in feminine singular, and an object relative clause with a head noun in neuter

---

2 Hybrid nouns denote a plurality but have the form of a singular noun, e.g., *djeca* ‘children’, *braća* ‘brothers’, etc. (see Corbett 1983a, 1991). They trigger plural agreement on the predicate:

(i) Djeca/            braća            plaču.  
children<sup>F.SG</sup>/    brothers<sup>F.SG</sup> cry<sup>3PL</sup>

singular. There were 54 fillers, which means there were 108 stimuli in total for each participant.

In the first experiment, there were two sessions for each participant. In the first session, participants were tested on agreement forms of the predicate with a subject containing conjoined nouns preceding the predicate. In the second session, the subject followed the predicate. These sessions were recorded over 15 days.

The second experiment was a written experiment conducted using a fill-in-the-blank questionnaire containing the same material as the oral experiment. Each example contained a model sentence and a stimulus – a conjoined phrase. Participants were instructed to write the form of the participle they felt was most appropriate with the conjoined phrase. As with the oral production experiment, the written experiment did not impose any time limits and was administered immediately after the oral experiment to the same participants who had participated in the oral experiment.

Nine sets of examples (illustrating nine conditions), each containing six sentences, plus 54 fillers, were presented. In the first session, the sentences contained a conjoined subject preceding the predicate, and in the second session, the order was reversed. This means that each participant was presented with 108 sentences in total.

Although we recognise that spoken language can better reflect native-speaker intuitions about language than can written language, as it does not allow for reflection on the *correctness* of an utterance, we conducted the written experiment to provide additional and clearer insight into agreement in BCS. We were aware of the fact that after being exposed to the examples in the oral experiment, the participants had already processed them when they encountered them for the second time in the written session. However, the written experiment was expected to contain fewer performance errors, since it gave the participants the possibility to skip examples or go back to them several times, which was not possible in the oral experiment.

There was a total of 6480 tokens resulting from nine lexicalisations of six conditions in two experiments (written and oral) and two sessions (pre-verbal vs. post-verbal placement of the conjoined subjects), completed by a total of 30 participants. However, the number of tokens actually considered in the analysis was 6074, as presented in Table 1 in the Appendix. Participants skipped some examples or provided unclear answers, so such examples were not taken into consideration. This means that for each condition (e.g., two masculine pre-verbal subjects in written elicitation), there were 180 tokens in total.

We present the results of our experiments documenting the existence of three distinct grammars of conjunct agreement in BCS: agreement with the highest conjunct, agreement with the closest conjunct, or agreement with the Boolean Phrase itself.

#### **4. UNIFORM GENDER CONJUNCTS**

We first investigated possible patterns of participial agreement with uniform gender conjuncts when both conjuncts are plural, and when uniform gender plural subjects occur both preverbally and postverbally. Such subjects largely elicit participial agreement

that corresponds to the gender of the two conjuncts. However, default masculine agreement occurs even when both conjuncts are feminine, or neuter, demonstrating that the ‘resolution rule’ of masculine agreement is attested even in uniform gender conjunctions (Corbett 1983a). However, this was registered predominantly in oral elicitation (44 examples oral, vs. 11 examples written).

#### 4.1 Preverbal Subjects

In oral elicitation, 24 examples (or 13.71%) of masculine default agreement were registered when both conjuncts were feminine, while only one example (0.56%) was registered in written elicitation, as in (3):<sup>3</sup>

- (3) *Fotografije i skice su prijavljeni na konkurs.* (6x)<sup>4</sup>  
 photos<sub>F</sub> and sketches<sub>F</sub> AUX registered<sub>M</sub> for competition  
 ‘Photographs and sketches were registered for the competition.’

In oral elicitation, 16 examples (8.89%) of masculine default agreement were registered when both conjuncts were neuter, as in (4):

- (4) *Poglavlja i uputstva su predani na čitanje.* (5x)  
 chapters<sub>N</sub> and instructions<sub>N</sub> AUX submitted<sub>M</sub> for reading  
 ‘Chapters and instructions were submitted for reading.’

There were five examples (2.79%) of masculine default agreement in written elicitation when both conjuncts were neuter.

Table 1: Numbers of participle agreement forms with uniform gender conjuncts with preverbal subjects

	WRITTEN			ORAL		
	M	F	N	M	F	N
M+M	179	0	0	178	0	0
F+F	1	179	0	24	151	0
N+N	5	0	174	16	0	164

3 It should be emphasised that all examples contained only inanimate subject NPs.

4 When the example was produced more than once, the number in parenthesis after the example indicates how many times that example was produced, e.g. the example in (3) was produced six times.

## 4.2 Postverbal Subjects

In postverbal elicitation, only nine examples of default masculine agreement were registered with the same gender conjuncts, all of them with feminine conjuncts, and these were almost equally distributed between oral (2.26%) and written elicitation (3.16%). Four examples were registered in oral elicitation, as in (5):

- (5) *Krajolikom su dominirali planine i rijeke.* (2x)  
 Landscape<sub>AUX</sub> dominated<sub>M</sub> mountains<sub>F</sub> and rivers<sub>F</sub>  
 ‘Mountains and rivers dominated the landscape.’

Five examples were registered in written elicitation.

Table 2: Numbers of participle agreement forms with uniform gender conjuncts with postverbal subjects

	WRITTEN			ORAL		
	M	F	N	M	F	N
M+M	180	0	0	180	0	0
F+F	5	153	0	4	173	0
N+N	0	1	176	0	1	175

## 5. MIXED GENDER CONJUNCTS

### 5.1 Preverbal Subjects

In written and oral elicitation with [Masc + Fem] plural conjuncts, both masculine and feminine agreement was present, but there were more examples of masculine agreement in written elicitation (150 examples = 83.80%) than in oral elicitation (128 = 73.56%), whereas there were more feminine agreement forms in oral elicitation (45 examples = 25.86%) than in written elicitation (27 = 15.08%). To summarise, the tendency for closest conjunct agreement (feminine) was more prominent in the oral than in the written experiments, although in both experiments, the masculine form of the participle was clearly dominant.

In written and oral elicitation with [Masc + Neut] plural conjuncts, both masculine and neuter agreement was present, but there were more examples of masculine agreement in the written elicitation (121 examples = 67.98%) than in oral elicitation (95 = 54.60%), whereas there were more neuter agreement forms in the oral elicitation (78 examples = 44.83%) than in the written elicitation (57 = 32.02%). Again, the tendency for closest conjunct agreement (neuter) was more prominent in the oral than in the written experiments, although in both experiments, the masculine form of the participle was dominant.

Table 3: Numbers of participle agreement forms with mixed gender (M+F and M+N) conjuncts: preverbal subjects

	WRITTEN			ORAL		
	M	F	N	M	F	N
M+F	150	27	2	128	45	1
M+N	121	0	57	95	1	78

In written and oral elicitation with [Fem + Masc] plural conjuncts, almost all examples with masculine agreement forms were elicited and with almost equal distribution in both the written (145 examples = 93.55%) and the oral elicitation (151 = 95.57%). However, there were 14 examples, nine (5.81%) in the written and five (3.16%) in the oral experiments, with feminine agreement, as in (6):

- (6) *Jedrilice i gliseri su uplovile u zaliv.* (6x) (written)  
 sailboats<sub>F</sub> and speedboats<sub>M</sub> AUX sailed<sub>F</sub> in bay  
 ‘Sailboats and speedboats sailed in the bay.’

Examples such as those in (6) contradict the claim by Bošković (2009) that highest conjunct agreement (HCA) in the preverbal position is not possible in BSC.

In written elicitation with [Fem + Neut] plural conjuncts, one half of elicited forms were default masculine (78 examples Masc (50%) vs. 58 Neut (37.18%). However, a considerable number of highest conjunct agreement forms (feminine) was recorded (20 examples = 12.82%), as in (7):

- (7) *Nagrade i priznanja su uručene na pozornici.* (5x)  
 prizes<sub>F</sub> and awards<sub>N</sub> AUX presented<sub>F</sub> on stage  
 ‘Prizes and awards were presented on the stage.’

In oral elicitation with [Fem + Neut] plural conjuncts, the majority of elicited forms were neuter agreement forms (87 examples Neut (53.37%) vs. 71 examples Masc (43.56%)). However, a small number of highest conjunct agreement forms (feminine) was recorded (5 examples = 3.07%).

Table 4: Numbers of participle agreement forms with mixed gender (F+M and F+N) conjuncts: preverbal subjects

	WRITTEN			ORAL		
	M	F	N	M	F	N
F+M	145	9	1	151	5	2
F+N	78	20	58	71	5	87

In written elicitation with [Neut + Masc] plural conjuncts, there were eight examples (5.23%) of highest conjunct agreement (neuter), as in (8), and all others were masculine (145 examples = 94.77%).

- (8) *Pitanja i problemi su napisana na tablu.* (2x)  
 questions<sub>N</sub> and problems<sub>M</sub> AUX written<sub>N</sub> on board  
 ‘Questions and problems were written on the board.’

In oral elicitation with [Neut + Masc] plural conjuncts, there were eleven examples (7.19%) of highest conjunct agreement (neuter), while others were masculine (141 examples = 92.16%).

In written elicitation with [Neut + Fem] plural conjuncts, all three agreement forms were recorded, and the same applies to oral elicitation. In written elicitation, the dominant forms were default masculine (96 examples = 54.55%); there were also 58 examples of feminine agreement (32.95%) and 22 examples (12.50%) of highest conjunct agreement (neuter), as in (9):

- (9) *Platna i skulpture su nestala u selidbi.* (5x)  
 canvases<sub>N</sub> and sculptures<sub>F</sub> AUX disappeared<sub>N</sub> in removal  
 ‘Canvases and sculptures were lost during removal.’

In oral elicitation with [Neut + Fem] plural conjuncts, the dominant forms were masculine (82 examples = 51.25%); there were 57 examples (35.63%) of feminine agreement and 21 examples (13.13%) of highest conjunct agreement (neuter).

Table 5: Numbers of participle agreement forms with mixed gender (N+M and N+F) conjuncts: preverbal subjects

	WRITTEN			ORAL		
	M	F	N	M	F	N
N+M	145	0	8	141	1	11
N+F	96	58	22	82	57	21

## 5.2 Postverbal Subjects

In written and oral elicitation with [Masc + Fem] plural conjuncts, only masculine agreement forms were elicited, with one exception in oral elicitation: a neuter agreement form, which is clearly a performance error. Similar results were obtained in written and oral elicitation with [Masc + Neut] plural conjuncts. In addition to masculine agreement forms, we recorded two examples with neuter participle forms in oral elicitation and three examples in written elicitation, as in (10):

- (10) *Na tržište su dospjela mirisovi i sjenila.* (2x) (oral)  
 to market AUX reached<sub>N</sub> perfumes<sub>M</sub> and eyeshadows<sub>N</sub>  
 ‘Perfumes and eyeshadows reached the market.’

Table 6: Numbers of participle agreement forms with mixed gender (M+F and M+N) conjuncts: postverbal subjects

	WRITTEN			ORAL		
	M	F	N	M	F	N
M+F	177	0	0	179	0	1
M+N	173	0	3	173	0	2

In written elicitation with [Fem + Masc] plural conjuncts, most forms were feminine (135 examples Fem = 88.82% vs. 15 Masc = 9.87%). In oral elicitation, there were 141 feminine forms (92.76%) and 11 examples of masculine forms (7.24%), as in (11):

- (11) *U roku su položeni vježbe i seminari.* (3x written, 1x oral)  
 in time AUX passed<sub>M</sub> exercises<sub>F</sub> and seminars<sub>M</sub>  
 ‘Exercises and seminars were passed on time.’

In written elicitation with [Fem + Neut] plural conjuncts, almost all forms were feminine. There were four examples (2.50%) of masculine (default) agreement in written elicitation and six examples (3.82%) in oral elicitation, as in (12):

- (12) *U dućan su stigli olovke i rumenila.* (oral)  
 in shop AUX delivered<sub>M</sub> pencils<sub>F</sub> and blushers<sub>N</sub>  
 ‘Pencils and blushers were delivered to the shop.’

In oral elicitation with [Fem + Neut] plural conjuncts, there were four neuter forms (2.55%), as in (13):

- (13) *Na petak su pomaknuta sjednice i vijeća.*  
 to Friday AUX rescheduled<sub>N</sub> meetings<sub>F</sub> and assemblies<sub>N</sub>  
 ‘Meetings and assemblies were rescheduled for Friday.’

Table 7: Numbers of participle agreement forms with mixed gender (F+M and F+N) conjuncts: postverbal subjects

	WRITTEN			ORAL		
	M	F	N	M	F	N
F+M	15	135	2	11	141	0
F+N	4	156	0	6	147	4



In written and oral elicitation with [Neut + Masc] plural conjuncts, only four masculine forms were registered – one (0.68%) in written and three (2.03%) in oral elicitation, as in (14):

- (14) *U vožnji su pomogli upozorenja i savjeti.* (2x) (oral)  
 in driving AUX assisted<sub>M</sub> warnings<sub>N</sub> and suggestions<sub>M</sub>  
 ‘Warnings and suggestions helped during ride.’

In written and oral elicitation with [Neut + Fem] plural conjuncts, most forms were neuter. There were nine examples with default masculine forms – three (1.68%) in written and six (3.45%) in oral elicitation, as in (15):

- (15) *U sobu su naručeni jaja i salate.* (written and oral)  
 to room AUX ordered<sub>M</sub> eggs<sub>N</sub> and salads<sub>F</sub>  
 ‘Eggs and salads were ordered to the room.’

However, there were ten feminine forms: four examples (2.23%) in written and six (3.45%) in oral elicitation, as in (16):

- (16) *Pred zoru su utihnule nevremena i poplave.* (5x) (oral)  
 before dawn AUX subsided<sub>F</sub> storms<sub>N</sub> and floods<sub>F</sub>  
 ‘Storms and floods subsided before dawn.’

Table 8: Numbers of participle agreement forms with mixed gender (N+M and N+F) conjuncts: postverbal subjects

	WRITTEN			ORAL		
	M	F	N	M	F	N
N+M	1	0	145	3	0	145
N+F	3	4	172	6	6	162

We will now discuss the examples of apparent agreement with the farthest conjunct in postverbal experiments. There were four such examples in the written experiments, all involving a neuter participle form with conjoined masculine and neuter plural nouns. Following the claim by Marušič et al. (2015) that a postverbal conjunction does not allow agreement with the second/ last/farthest conjunct, we treated this type of agreement as performance errors, rather than a separate agreement strategy, as in (17):

- (17) *U sali su operisana zglobovi i stopala.*  
 in hall AUX operated<sub>N</sub> joints<sub>M</sub> and feet<sub>N</sub>  
 ‘Joints and feet were operated in the hall.’

There were also eleven such examples in the oral experiments. Five examples involve a neuter participle form with conjoined feminine and neuter plural nouns, as in (18):

- (18) *Na petak su pomaknuta sjednice i vijeća.*  
to Friday AUX rescheduled<sub>N</sub> meetings<sub>F</sub> and assemblies<sub>N</sub>  
'Meetings and assemblies were rescheduled for Friday.'

Six examples involve a feminine participle form with conjoined neuter and feminine plural nouns, as in (19):

- (19) *U sobu su naručene jaja i salate.*  
to room AUX ordered<sub>F</sub> eggs<sub>N</sub> and salads<sub>F</sub>  
'Eggs and salads were ordered to the room.'

## 6. DISCUSSION

The results of our experiments reveal that the default masculine agreement is more prominent when both conjuncts are not of the same gender, but rather of mixed gender, especially feminine and neuter. In the preverbal written elicitation with [Fem + Neut] conjuncts, one half of participle forms were masculine (78 = 50%); there were 58 neuter forms (37.18%) and 20 feminine forms (12.82%). This clearly demonstrates that participants in our experiment used three strategies of agreement: agreement with the maximal projection – a Boolean Phrase (78 masculine participle forms), agreement with the nearest conjunct (58 neuter participle forms), and agreement with the conjunct which is hierarchically the highest conjunct (20 feminine forms). In the preverbal written elicitation with [Neut + Fem] conjuncts, again the majority of participle forms were masculine (96 = 54.55%); there were 58 feminine forms (32.95%) and 22 neuter forms (12.50%). Very similar results were obtained in preverbal oral elicitation with [Neut + Fem] conjuncts: 82 examples (51.25%) with masculine agreement, 57 examples (35.63%) of feminine agreement, and 21 examples (13.13%) of agreement with the highest, neuter conjunct. However, in preverbal oral elicitation with [Fem + Neut] conjuncts, there were fewer examples of agreement with the highest, feminine conjunct (only 5 examples = 3.07%), with most examples of agreement with the nearest, neuter conjunct (87 examples = 53.37%), and 71 examples (43.56%) of default masculine agreement.

The total number of elicited examples, both written and oral, with [Neut + Fem] and [Fem + Neut] conjuncts in preverbal position was 655. The dominant form of agreement was the default masculine agreement (327 examples, or (roughly) 50%), then the agreement with the closest conjunct (260 examples, or (roughly) 40%), and the least represented was the agreement with the highest conjunct (68 examples, or (roughly) 10%). The percentage of 10% is not small, so these examples cannot be treated as performance errors, and therefore we conclude that agreement with the highest conjunct is the third strategy of agreement used by native speakers. This contradicts the claim by Bošković (2009) that highest conjunct agreement in the preverbal position is not possible in BCS.

Then we investigated possible patterns of participial agreement with [Fem + Neut] and [Neut + Fem] conjuncts in the postverbal position. In the postverbal written elicitation with [Fem + Neut] conjuncts, there were no examples of agreement with the furthest, neuter conjunct. There were only four examples (2.50%) of default, masculine agreement, and 156 examples (97.50%) of nearest, feminine conjunct agreement. In postverbal written elicitation with [Neut + Fem] conjuncts, the majority of examples (172) were again examples of agreement with the nearest, neuter conjunct; there were three examples (1.68%) of default, masculine agreement, and four examples (2.23%) of agreement with the furthest, feminine conjunct. In postverbal oral elicitation with [Fem + Neut] conjuncts, there were 147 examples of agreement with the nearest, feminine conjunct, six examples of default, masculine agreement, and four examples (2.55%) of agreement with the furthest, neuter conjunct. In postverbal oral elicitation with [Neut + Fem] conjuncts, there were 162 examples of agreement with the nearest, neuter conjunct, six examples of default, masculine agreement, and six examples (3.45%) of agreement with the furthest, feminine conjunct.

The total number of elicited examples, both written and oral, with [Neut + Fem] and [Fem + Neut] conjuncts in postverbal position was 670. The dominant form of agreement was the agreement with the nearest conjunct (637 examples, or (roughly) 95%), then the default, masculine agreement (19 examples, or (roughly) 3%), and the least represented was the agreement with the furthest conjunct (14 examples, or (roughly) 2%). The percentage of 2% is too small; therefore, these examples should be treated as performance errors rather than a separate agreement strategy (agreement with the furthest conjunct). Thus, our results confirm the claim by Marušič et al. (2015) that a postverbal conjunction does not allow agreement with the second/last/farthest conjunct. The summary of results is presented in the tables in the Appendix.

Finally, we observed the following differences between the results of the written and the oral experiments. With uniform gender conjuncts, there were considerably more default agreement forms in the oral (40 examples) than in the written experiments (6 examples), but only with preverbal subjects. On the other hand, with postverbal subjects, almost the same number of default forms was registered in both the oral (4 examples) and the written experiments (5 examples).

With mixed gender conjuncts in the preverbal position, there were more default agreement forms in the written (735 examples) than in the oral experiments (668 examples). However, there were more examples of closest conjunct agreement in the oral (267 examples) than in the written experiments (200 examples). Finally, there were more examples of highest conjunct agreement in the written (59 examples) than in the oral experiments (42 examples).

With mixed gender conjuncts in the postverbal position, closest conjunct agreement is a dominant form of agreement; it was almost equally represented both in the written (958 examples) and the oral experiments (947 examples). Also, examples of default agreement were almost equally distributed: 23 examples in the written and 26 examples in the oral experiments.

## 7. CONCLUSION

In this paper, we presented the results of experiments which tested the strategies of subject-predicate agreement in BCS, where the subject consists of conjoined noun phrases. Our experiments confirmed that agreement strategies attested by Marušič et al. (2015) for Slovenian exist in BCS as well. Consequently, BCS, like Slovenian, has three distinct strategies of agreement: 1. agreement with the maximal projection – a Boolean Phrase (&P), 2. agreement with the conjunct which is closest to the participle, and 3. agreement with the conjunct which is hierarchically the highest. The results obtained in our experiment justify the claim that Highest Conjunct Agreement (HCA) in contexts with preverbal subjects should be treated as a legitimate agreement strategy in BCS, as opposed to agreement with the furthest conjunct in contexts with postverbal subjects, which we claim to be the result of performance errors. These results are contrary to Bošković's findings (2009), in which he does not acknowledge HCA as a legitimate strategy, however, our results do confirm the findings of Marušič et al. (2015).

## References

- AOUN, Joseph/Elabbas BENMAMOUN/Dominique SPORTICHE (1999) "Further Remarks on First Conjunct Agreement." *Linguistic Inquiry* 30, 669–681.
- BENMAMOUN, Elabbas/Archana BHATIA/Maria POLINSKY (2010) "Closest Conjunct Agreement in Head Final Languages." In: J. van Craenenbroeck (ed.), *Linguistic Variation Yearbook 2009*, 67–88.
- BHATT, Rajesh/Martin WALKOW (2013) "Locating Agreement in Grammar: An Argument from Agreement in Conjunctions." *Natural Language and Linguistic Theory* 31/4, 951–1013.
- BOCK, Kathryn/Carol MILLER (1991) "Broken Agreement." *Cognitive Psychology* 23, 45–93.
- BOŠKOVIĆ, Željko (2009) "Unifying First and Last Conjunct Agreement." *Natural Language and Linguistic Theory* 27, 455–496.
- BOŠKOVIĆ, Željko (2010) "Conjunct-Sensitive Agreement: Serbo-Croatian vs Russian." In: G. Zybatow/P. Dudchuk/S. Minor/E. Pschehotskaya (eds), *Formal Studies in Slavic Linguistics: Proceedings of Formal Description of Slavic Languages 7.5*. Frankfurt am Main: Peter Lang, 31–48.
- CORBETT, Greville (1983a) "Resolution Rules: Agreement in Person, Number and Gender." In: G. Gazdar/E. Klein/G. Pullum (eds), *Order, Concord and Constituency*. Dordrecht: Foris, 175–214.
- CORBETT, Greville (1983b) *Hierarchies, Targets, and Controllers: Agreement Patterns in Slavic*. London: Croom Helm.
- CORBETT, Greville (1991) *Gender*. Cambridge: Cambridge University Press.
- CULICOVER, Peter/Ray JACKENDOFF (2005) *Simpler Syntax*. Oxford: Oxford University Press.

- DRUMMOND, Alex (2011) *IbexFarm*. Version 0.2.7. Available <http://spellout.net/ibexfarm>
- FRANCK, Julie (2011) “Reaching Agreement as a Core Syntactic Process. Commentary on Bock and Middelton ‘Reaching Agreement’.” *Natural Language and Linguistic Theory* 29, 1071–1086.
- FRANCK, Julie/Glenda LASSI/ Ulrich FRAUENFELDER/Luigi RIZZI (2006) “Agreement and Movement: A Syntactic Analysis of Attraction.” *Cognition* 101, 173–216.
- FRANCK, Julie/ Ulrich FRAUENFELDER/Luigi RIZZI (2007) “A Syntactic Analysis of Interference in Subject-Verb Agreement.” *MIT Working Papers in Linguistics* 53, 173–190.
- MARUŠIČ, Franc/Andrew NEVINS/Amanda SAKSIDA (2007) “Last-Conjunct Agreement in Slovenian.” In: R. Compton/M. Goledzinowska/U. Savchenko (eds), *Formal Approaches to Slavic Linguistics* 15. Ann Arbor, Mich.: Michigan Slavic Publications, 210–227.
- MARUŠIČ, Lanko/Andrew NEVINS/William BADECKER (2015) “The Grammars of Conjunction Agreement in Slovenian.” *Syntax* 18/1, 39–77.
- MUNN, Alan (1999) “First Conjunct Agreement: Against a Clausal Analysis.” *Linguistic Inquiry* 30/4, 643–668.
- PUŠKAR, Zorica/Andrew MURPHY (2015) “Closest Conjunct Agreement in Serbo-Croatian: A Rule-Ordering Account.” In: A. Assmann/S. Bank/D. Georgi/T. Klein/P. Weisser/E. Zimmermann (eds), *Topics at Infl. Vol 92 of Linguistische Arbeitsberichte* (LAB), Universität Leipzig, 441–482.

## Appendix 1.

Table 9: Numbers of participle agreement forms with uniform and mixed gender conjuncts

	SV						VS					
	WRITTEN			ORAL			WRITTEN			ORAL		
	M	F	N	M	F	N	M	F	N	M	F	N
<b>M+M</b>	179	0	0	178	0	0	180	0	0	180	0	0
<b>F+F</b>	1	179	0	24	151	0	5	153	0	4	173	0
<b>N+N</b>	5	0	174	16	0	164	0	1	176	0	1	175
<b>M+F</b>	150	27	2	128	45	1	177	0	0	179	0	1
<b>M+N</b>	121	0	57	95	1	78	173	0	3	173	0	2
<b>F+M</b>	145	9	1	151	5	2	15	135	2	11	141	0
<b>F+N</b>	78	20	58	71	5	87	4	156	0	6	147	4
<b>N+M</b>	145	0	8	141	1	11	1	0	145	3	0	145
<b>N+F</b>	96	58	22	82	57	21	3	4	172	6	6	162

Table 10: Numbers of performance errors, highest-conjunct agreement, and furthest-conjunct agreement

	SUBJECT PREVERBAL		SUBJECT POSTVERBAL	
	WRITTEN	ORAL	WRITTEN	ORAL
M+M				
F+F				
N+N			1F (PE)	1F (PE)
M+F	2N (PE)	1N (PE)		1N (PE)
M+N		1F (PE)	4N (FC)	1F (PE)
F+M	6F (HC); 1N(PE)	3F(HC); 1N(PE)	2N (PE)	1MSg (PE)
F+N	19F (HC)	4F (HC)		5N (FC)
N+M	10N (HC); 1F(PE)	7N(HC)		
N+F	22N (HC)	6N (HC)	1FSg; 1NSg (PE)	6F (FC), 5 (PE)

PE = performance error

HC = highest conjunct agreement

FC = furthest conjunct agreement

Table 11: Percentages of participle agreement forms with uniform and mixed-gender conjuncts

	Subject preverbal						Subject postverbal					
	Written			Oral			Written			Oral		
	M	F	N	M	F	N	M	F	N	M	F	N
M+M	100%	0	0	100%	0	0	100%	0	0	100%	0	0
F+F	0.56	99.44	0	13.71	86.29	0	3.16	96.84	0	2.26	97.74	0
N+N	2.79	0	97.21	8.89	0	91.11	0	0.56	99.44	0	0.57	99.43
M+F	83.80	15.08	1.12	73.56	25.86	0.57	100%	0	0	99.44	0	0.56
M+N	67.98	0	32.02	54.60	0.57	44.83	98.30	0	1.70	98.86	0	1.14
F+M	93.55	5.81	0.65	95.57	3.16	1.27	9.87	88.82	1.32	7.24	92.76	0
F+N	50.00	12.82	37.18	43.56	3.07	53.37	2.50	97.50	0	3.82	93.63	2.55
N+M	94.77	0	5.23	92.16	0.65	7.19	0.68	0	99.32	2.03	0	97.97
N+F	54.55	32.95	12.50	51.25	35.63	13.13	1.68	2.23	96.09	3.45	3.45	93.10

### Summary

#### STRATEGIES OF PARTICIPLE AGREEMENT WITH CONJOINED SUBJECTS IN BOSNIAN/CROATIAN/SERBIAN

In this paper, we demonstrate that Bosnian/Croatian/Serbian (BCS), like Slovenian, has three distinct strategies of subject-predicate agreement when the subject consists of conjoined noun phrases: 1. agreement with the maximal projection – a Boolean Phrase (&P), 2. agreement with the conjunct which is closest to the participle, and 3. agreement with the conjunct which is hierarchically the highest. In order to test the initial hypothesis that there are three agreement strategies, we conducted a controlled experimental study of the morphosyntax of agreement between conjoined subjects and participles in BCS, which consisted of an oral production experiment and a written production experiment. These experiments revealed a high presence of default agreement and closest conjunct agreement in the language. 50% of preverbal conjoined phrases elicited the default masculine agreement and 95% of postverbal conjoined noun phrases elicited the closest conjunct agreement. However, the bulk of the analysis focused on the possibility of treating the highest conjunct agreement (HCA) as a legitimate agreement strategy. 7% of all of the agreement forms in the subject preverbal (SV) examples demonstrated HCA. These figures increased to 13% if individual conditions

were considered. Last conjunct agreement (LCA) for subject postverbal (VS) examples, on the other hand, was only present in 1% of the examples. For this reason, we classified them as performance errors and refuted LCA as an agreement strategy. These results are contrary to Bošković's findings (2009), in which he does not acknowledge HCA as a legitimate strategy, however, our results do confirm the findings of Marušič et al. (2015).

**Keywords:** conjunction, closest-conjunct agreement, first-conjunct agreement, experimental syntax

Povzetek  
STRATEGIJE UJEMANJA DELEŽNIKA S KOORDINIRANIMI OSEBKI  
V BOSANSKEM/HRVAŠKEM/SRBSKEM JEZIKU

V članku pokažemo, da ima bosanski/hrvaški/srbski jezik (BHS) tako kot slovenščina tri strategije ujemanja med osebkom in predikatom, ko je osebek koordinirana samostalniška zveza: 1. ujemanje z maksimalno projekcijo – Boolejeva zveza (&P); 2. ujemanje s koordinirano prvino, ki je bližja deležniku; 3. ujemanje s koordinirano prvino, ki je najvišje v hierarhiji. Da bi preverili podmeno o treh strategijah ujemanja, smo izvedli kontrolirano eksperimentalno študijo oblikoslovno-skladenjskega ujemanja med koordiniranimi osebki in deležniki v BHS, ki je vključevala eksperimenta z govornim in pisnim jezikom. Eksperimenta sta pokazala visoko prisotnost privzetega ujemanja in ujemanja z najbližjo koordinirano prvino v osebku. Petdeset odstotkov predglagolskih koordiniranih zvez je pokazalo privzeto ujemanje z moško obliko, petindevetdeset odstotkov zaglagolskih koordiniranih samostalniških zvez pa se je ujemalo po načelu najbližje koordinirane prvine. Analiza je bila v največji meri osredinjena na ustreznost ujemanja s koordinirano prvino, ki je najvišje v hierarhiji (HCA). Slednja se je pojavila v sedmih odstotkih oblik, kjer je bil osebek v predglagolskem položaju. Odstotek naraste na 13, če upoštevamo posamezne pogoje. Ujemanje z zadnjo koordinirano prvino pri zaglagolskih osebkih (LCA) pa je bilo prisotno le v enem odstotku primerov, zato smo jih opredelili kot napake v rabi jezika in zavrnili LCA kot ustrezno strategijo ujemanja. Rezultati nasprotujejo študiji Boškovića (2009), ki zavrača HCA kot ustrezno strategijo, potrjujejo pa izsledke Marušiča et al. (2015).

**Ključne besede:** konjunkcija, ujemanje z najbližjo koordinirano prvino, ujemanje s prvo koordinirano prvino, eksperimentalna skladnja







## IN THE SOUTH SLAVONIC GARDEN: LANDSCAPING THE LANDSCAPE OF ARGUMENTS AND NON-ARGUMENTS\*\*

### 1. INTRODUCTION

Over the last couple of decades, the phenomenon of Cognate Object Construction (COC) has been extensively discussed in theoretical linguistics. Though widely discussed, even the status of the relevant nominals remains a matter of controversy. Researchers disagree not only on the syntactic and semantic status of what is referred to as Cognate Objects (COs), but also on the judgments regarding them:

- (1) a) \**An uneventful life was lived by Harry.* (Jones 1988: 91)  
b) *A good life was lived by Susan.* (Rice 1988: 210)

In this paper, I address syntactic, semantic, and pragmatic puzzles that COs raise and also demonstrate that judgments as in (1) are not accidental or idiolectal. Rather, when viewed through the prism of languages under consideration here, they become congruent and can straightforwardly be explained.

The core of my analysis is the following. Based on the data from Serbo-Croatian (SC) and Slovene (SLO), I identify two types of COs: i) arguments (ACOs) & ii) predicates (non-ACOs). The latter are first order predicates, while the former are arguments in the neo-Davidsonian sense. The notion of “cognatehood” is irrelevant in syntactic terms for either group, though it becomes relevant for the latter in a pragmatic sense. Syntactically, the non-ACOs are adjuncts; the fact that they cross-linguistically occur with unaccusatives is unproblematic. As for ACOs, I argue that verbs with which they appear are regular transitive and regular unergative verbs (see Marelj 2015).

The paper is organised as follows: In section 2, I present core issues that theories on COs must address. In section 3, based on SC and SLO data, I discuss CO-tests. The results of this section lead me to identify two types of COs, while in section 4 I tackle non-ACOs and in section 5 I tackle ACOs. Finally, in section 6, I present the conclusion of this paper.

---

\* m.marelj@uu.nl

\*\* I am grateful to the three anonymous reviewers of *Linguistica* for their valuable input, Marko Hladnik for the Slovene and Ora Matushansky for the Russia data. This research was partially supported by the *Aspasia* grant, awarded to me by the *Netherlands Organisation for Scientific Research (NWO)*, which I gratefully acknowledge here. Any errors that remain are my sole responsibility.

## 2. MYSTERIOUS MYSTERIOUSNESS

Crosslinguistically, it is examples like (2) – featuring intransitive verbs accompanied by etymologically related nouns – that are typically found in the literature illustrating COC. The morphological relatedness is not the only characteristic of the classical cases of COs. Though there is a well-known generalisation that COs appear only with unergatives (see Hale and Keyser 2002 a.o.), it is the unaccusative *die* as in (3) that often crops up at the most typical example of COC.

- (2) a) *He dreams a dream.*  
 b) *Živi život.* [SC]  
 lives.3sg life  
 ‘He lives a life.’
- (3) a) *He died a natural death.*  
 b) *Umro je prirodnom smrću.* [SC]  
 die.prt.m.sg AUX.3sg natural death  
 ‘He died a natural death.’

### 2.1 Scarsity of Syntactic Space

The well-formed nature of examples such as (3) raises the question of the scarcity of syntactic space. Since the sole syntactic argument of unaccusatives (4b) begins its life as a complement of a verb, it is unclear how, to start with, “an additional complement” would enter the structure.

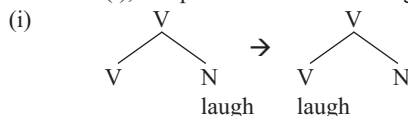
- (4) a) *He* [<sub>VP/VP</sub> *he-dreamt*] [Unergative]  
 b) *He* [<sub>VP</sub> [<sub>V</sub> *died he*]] [Unaccusative]

The fact that unergatives and unaccusatives are quite different is uncontroversial. One of the tests that seems to be cross-linguistically valid pertains to the fact that the past participial can be used attributively only with the internal arguments of relevant verbs. This prediction is borne out in (5).

- (5) a) *the melted snow/the departed guest*  
 b) \**the shouted child*

Slavonic languages behave *on a par* with English. For instance, a past participle in SC (active and passive alike) modifies internal arguments only (6):<sup>1</sup>

1 Under an analysis such as (4a), the question of syntactic space does not seem to arise for unergatives. However, considering the fact that the prevalent view of unergatives is rooted in Hale and Keyser’s work (i), this problem arises for unergatives as well. I return to this issue in 5.2.



- (6) a) \**plesani čovek*  
 danced man  
 b) *razbijeni/pali*                      *andjeo*  
 broken/fallen.prt.m.sg    angel  
 ‘a broken/fallen angel’

## 2.2 What’s in a Modifier?

COs do not differ only in their syntax. *Smrt* (death) cannot possibly occur unmodified:

- (7) #/\**Umrli*        *je*                      *smrti*    [SLO]  
 die.prt.m.sg    AUX.3sg    death

This behaviour is neither restricted to *umreti* (die) nor to SLO. However, the obligatory nature of modification does not hold in all instances of COs. This curious behaviour has been left unexplained.

## 3. LANDSCAPING THE LANDSCAPE...

As the data from SC and SLO are intended to be used as a sort of “litmus” test for the deliberation on the status of the data like (1), it is only prudent to start by submitting the relevant data to the standard tests on COs.

### 3.1 Classifying Cognate Objects

Examples in (8) – (11) represent the most canonical tests on COs. As “die” and “dream” are typically found in the literature on English, below I use their equivalents in SC and SLO as well.

- (8) Passivisation:
- a) *San*                      *je*                      *sanjan.*    [SC]  
 Dream AUX.3sg dreamt.prt.m.sg  
 ‘The dream was dreamt.’
- b) \**Užasna smrt*        *je*                      *umrna*  
 horrible death    AUX.3sg died.prt.f.sg
- c) *Sanje*                      *so*                      *sanjane.*    [SLO]  
 dream.pl    AUX.3pl    dreamt.prt.f.pl  
 ‘The dreams were dreamt.’
- d) \**Strašna smrt*        *je*                      *umrta*  
 horrible death    AUX.3sg    died.prt.f.sg

(9) Pronominalisation:

a) *Svake noći je sanjao zastrašujućisan.* [SC]  
every night AUX.3sg dreamt.prt.m.sg terrifying dream  
'Every night he dreamt a terrifying dream.'

a') *Svake noći ga je sanjao.*  
every night CL.acc.m AUX.3sg dreamt  
'Every night he dreamt it.'

b) *Svake noći je umirao užasnom smrću.*  
Every night AUX died horrible death  
'Every night he died a horrible death.'

b') \**Svake noći je njom umirao.*  
Every night AUX.3sg CL.instr.f died.prt.m.sg.

c) *Vsako noć je sanjal strašne sanje.* [SLO]  
every night AUX.3sg dreamt.prt.m.sg terrifying dream  
'Every night he dreamt a terrifying dream.'

c') *Vsako noć jih je sanjal.*  
every night CL.acc.f.pl AUX.3sg dreamt  
'Every night he dreamt them.'

d) *Vsako noć je umrl strašne smrti.*  
Every night AUX died.prt.m.sg horrible death  
'Every night he died a horrible death.'

d') \**Vsako noć jo je umrl.*  
Every night AUX.3sg CL.gen.f died.prt.m.sg.

(10) Definiteness restriction:

a) *Sanjao je taj/ovaj zastrašujući san.* [SC]  
dreamt AUX.3sg this/that terrifying.acc dream.acc  
'He dreamt that/this terrifying dream.'

b) \**Umro je tom/ovom užasnom smrću.*  
die.prt.m.sg AUX.3sg this/that horrible.instr death.instr

c) *Sanjal je te/tiste strašne sanje.* [SLO]  
dreamt.prt.m.sg AUX.3sg this/that terrifying.acc dream.acc  
*He dreamt this/that terrifying dream.*

d) \**Umrl je te/tste strašne smrti.*  
 die.prt.m.sg AUX.3sg this/that horrible.gen death.gen

(11) Can be questioned:<sup>2</sup>

a) *Šta je sanjao?* [SC]  
 what AUX.3sg dream.prt.m.sg  
 ‘What did he dream?’

b) \**Šta je umro?*  
 What.ACC AUX.3SG died.prt.m.sg

c) *Kaj je sanjal?* [SLO]  
 what AUX.3sg dream.prt.m.sg  
 ‘What did he dream?’

d) \**Kaj je umrl?*  
 what.acc AUX.3sg died.prt.m.sg

Based on (8) – (11), the behaviour of *sanjati* (dream) i *umreti* (die) in SC and SLO seems consistent with the behaviour of their English equivalents. The data seem to require a sort of a “hybrid” approach, since some of the COs behave like run-of-the-mill arguments, while others behave like non-arguments/adjuncts. Before I address the two types of COs in SC and SLO, however, I must first address the relevance of the morphological case.

### 3.2 What’s in a Case?

Though (8)–(11) seem to indicate that English, on the one hand, and SC and SLO, on the other hand, pattern fully alike, this is not entirely true. The data in these morphologically robust languages seem to give rise to another generalisation; the “real” arguments appear in ACC, while the non-arguments appear in INSTR (SC) and GEN (SLO). At first blush, then, it seems that we have an additional morphosyntactic test to differentiate the argumental from the non-argumental instances of COs, thus having another piece of evidence to support that the morphological structure reflects the underlying structure. For instance, it is typically argued that non-verbal predicates bear INSTR as a predicate case (see Pereltsvaig 1999 for the discussion and references on Russian) (12).

2 I kindly ask (native) speakers of SC and SLO to grant me a bit of patience. Though they might exclaim at this point that what happens in (11b) and (11d) is an instance of a case mismatch that goes beyond the issues at stake here, I would like to remind them that leaving such examples in the text here underscores the relevance of exploring Slavonic languages like SC and SLO, since the traditional generalisation about COC in the literature (see Jones 1988, for instance) is that languages that express the morphological case overtly, such as Arabic, German, or Latin, select ACC, “a semantically empty Case to satisfy morphological requirements” (see Jones 1988: 109, but also Moltman 1989, for instance).

- (12) *On stal učitelem* [Russian]  
 He become.prt.m.sg teacher.instr  
 ‘He became a teacher.’

I argue here that a correlation between some “designated” predicate case and non-verbal predicates in South Slavonic simply does not hold. Contrary to standard views in the literature on Slavonic, I argue that there is no deeper relevance or meaning to the case marking here. The overt morphology here is simply idiosyncratic. Trivially, to begin with (quite like in other Slavonic languages), instrumental is not restricted to non-verbal predicates. For instance, verbs such as *vladati* (rule), *upravljati* (manage), *dominirati* (dominate) and *trajati* (last) in SC have INSTR internal arguments.

Not only is the correlation not two-directional, but, more importantly, it is not one-directional either. Though there are a couple of exceptions as in (13), generally, non-verbal predicates in SC occur as NOM, not INSTR.

- (13) *Postao je učitelj/učiteljem*  
 become.prt.m.sg AUX. teacher.nom/teacher.instr  
 ‘He became a teacher.’

Slovene underscores this point about the lack of the correlation in this respect even more strongly. Firstly, the canonical cases of non-arguments under consideration here in Slovene do not even consistently appear in the same “alleged” predicate case. Rather, some are genitive (14a), while others are prepositional instrumental (14b).

- (14) a) *Umrl je naravne smrti.* [SLO]  
 die.prt.m.sg AUX.3sg natural death.gen  
 ‘He died a natural death.’  
 b) *Rasel je z neustavljivo rastjo.*<sup>3</sup>  
 grew.prt.m.sg AUX.3sg with unstoppable growth.instr  
 ‘He grew the unstoppable growth.’

Moreover, the cases of INSTR non-verbal predicates like Russian (12) or even in SC (13) are always nominative.<sup>4</sup>

3 Cross-linguistically, it seems that the data on unaccusatives are met with varying reactions from native speakers. It is not obvious to me why this is so. Note that this is also the case with (14b), where R1 and my native speaker informant differ.

4 Needless to say, quite like SC, SLO also has “real” complements in genitive (ia) and prepositional instrumental case (ib):

- (i) a) *Boji se Petra.* [SLO]  
 fear.3sg CL Peter.gen  
 ‘He fears Peter.’  
 b) *Upravljaj je s tovarno.*  
 Managed prt.m.sg AUX.3sg with factory.instr  
 ‘He managed a factory.’

- (15) *Postal je učitelj.* [SLO]  
 became AUX.3sg teacher.nom  
 ‘He became a teacher.’

#### 4 ZOOMING IN ON NON-ACOS

##### 4.1 Where Adverbs and Cognates Meet

As any native speaker of SC and SLO can confirm (see also fn. 3), the ungrammaticality of (11b) and (11d) seems far more trivial than that of the English (16). INSTR- and GEN-nominals cannot be questioned using ACC *wh*-word *šta/kaj* (what.ACC) and *wh*-word *čime* (what.INSTR)/*česa* are required.

- (16) \**What did he die?* (Massam 1990:164)

Note importantly, however, that the use of the morphologically “appropriate” *wh*-word will not improve the grammaticality of SLO and SC examples (17).

- (17) a) \**Čime je umro? Užasnom smrću.* [SC]  
 what.instr AUX.3sg die.prt.m.sg terrible.instr death.instr  
 b) \**Česa je umrl? Strašne smrti.* [SLO]  
 what.gen AUX.3SG die.prt.m.sg horrible.gen death.gen

The way to question the relevant SC and SLO examples is by using the *wh*-word “how”:

- (18) *Kako je umro? Užasnom smrću.* [SC]  
 How AUX.3sg died.prt.m.sg horrible death.instr  
 ‘How did he die? A horrible death.’

What does the use of this *wh*-element tell us about the syntax and semantics of non-ACOs? *Wh*-word *kako* (how) appears with adverbial modification – specifically, with manner adverbs. Consequently, interpretatively, non-ACOs seem to be comparable to manner adverbs (19).

- (19) *Kako se ponaša? Pristojno/loše...* [SC]  
 how CL behave.3sg decently/badly  
 ‘How does he behave? Decently/badly....’

The fact that examples such as (20a) can be paraphrased using sentences like (20b) further underscores the parallelism:



- (20) a) *Smejao se grohotnim smehom.* [SC]  
 laugh.prt.m.sg CL loud.instr laugher.instr  
 ‘He laughed a loud laughter.’
- b) *Grohotno se smejao.*  
 loudly CL laugh.prt.m.sg  
 ‘He laughed loudly.’

The question that now arises is the following: How should this parallelism be understood and formalised?

#### 4.2 On Adverbs and Analogies

Since we have noted the parallelism between manner adverbs and non-ACOs, we can explore whether or not an approach to the latter can be modelled based on the established approach to the former.

The most elegant approach to manner adverbs originates in Davidson (1967). In Davidson’s original proposal, the event argument is accommodated by analysing the predicate as having one more argument place than is assumed in traditional analyses. The event argument (e) is existentially quantified, with the result that a sentence like (21a) takes on a logical structure, such as in (21b). Manner adverbs are added conjunctively and predicate the event argument directly, as in (21c).

- (21) a) *Tristram ate a snowflake (carefully).*  
 b)  $\exists e$  eat (Tristram, a snowflake, e)  
 c)  $\exists e$  [EAT (e, T, s) & carefully (e)]

The neo-Davidsonian tradition (Higginbotham 1987, Parsons 1990) deviates from the original arguments into separate conjuncts as well:

- (22)  $\exists e$  [ working (e) & Agent (e)=T & Theme (e)=s & carefully (e)]

The status of event argument modifiers seems to be exactly the right “description” of what non-ACOs in SC and SLO do. Indeed, my hypothesis is that the non-ACOs in SC and SLO are event predicates.<sup>5</sup> Let us explore this hypothesis further.

#### 4.3 Referentiality, or One More Look at the Pronominalisation Test

Predicates are non-referential. For nominal predicates, this typically means that they are indefinite (see Higginbotham 1987 a.o.). Canonical cases that illustrate this include the bare NPs in languages like Dutch, which appear in predicative, but are barred from the argument positions. Though using bare NPs will not be a test in SC

<sup>5</sup> See Moltmann (1989) for the same conclusion regarding English and German COs of verbs like “die”.

and SLO, using pronominalisation as a test is a good way to establish the semantic status of a nominal.<sup>6</sup>

The rationale is straightforward: pronouns canonically refer to individuals and not to properties. Arguments always have referents and, consequently, can be pronominalised. Predicates do not have referents and hence cannot be pronominalised (23).

- (23) a) *Postao je matematičar/\*on* [primary predicate]  
 became.prt.m.sg AUX mathematician/he  
 ‘He became a mathematician/\*he’
- b) *Smatraju Mariju lepoticom/\*njome* [secondary predicate]  
 consider.3.pl Marija.acc beauty.instr/her.instr  
 ‘They consider Maria a beauty/\*her’

Note now that the non-ACOs behave on a par with the predicates (24) and quite unlike the regular arguments (25) with respect to pronominalisation (recall also (9) please) and (impersonal) passivisation (recall also (8) please).

- (24) a) *Nasmejao se grohotnim smehom/\*njim.* [INSTR-non-ACO]  
 laughed.prt.m.sg CL loud laughter.instr/him.instr  
 ‘He laughed a loud laughter/\*it.’
- b) *\*Grohotnim smehom je/se nasmejalo.*  
 loud laughter.instr AUX/CL laughed.prt.n.sg
- (25) a) *On upravlja fabrikom/njom.* [INSTR-argument]  
 he manage.3SG factory/it-INSTR  
 ‘He manages a factory/it.’
- b) *Fabrikom je upravljano (od strane radnika).*  
 factory.instr AUX managed prt.n.sg (by the workers)  
 ‘The factory was managed by the workers.’

In what follows, I demonstrate how two important properties of non-ACOs (see section 2) follow from the simple hypothesis that non-ACOs are event predicates.

#### 4.4 Mystery of Modification

The modification of non-ACOs need not always be appositive (26), but must always be present. The data below are from SC, but this is cross-linguistically true.

6 There is a known caveat here; it has been long established that there are pronouns the denotation of which is a property, rather than an individual (see Jespersen 1927; Williams 1983, a.o.), but [+human] pronouns are restricted to individuals.

- (26) *Umro je smrću mučenika.*  
 die.prt.m.3sg AUX.3sg death.instr martyr.gen  
 ‘He died a martyr’s death’

Though sentences like (7) are typically marked as ungrammatical in the literature, my assumption is that there is nothing wrong with their structure. I argue that they are pragmatically odd and that this oddness is the result of an interplay between their semantic status as predicates and their cognatehood.

The relevant questions are the following:

- a) Is it relevant for grammar as such that non-ACOs are etymologically related to relevant verbs or is this relatedness an idiosyncratic feature with no relevance whatsoever?
- b) Why is the modification of ACOs obligatory?

My hypothesis is that these two questions are directly related. Namely, I only partially agree with Perelsvaig’s (Perelsvaig 1999) conclusion that cognatehood is completely and fully irrelevant and inconsequential. Specifically, I concur with her conclusion that relatedness of relevant nominals is irrelevant in the cases of ACOs, the structural position and the denotation of which are no different than those of run-of-the-mill complements unrelated to the relevant verb. I argue, however, that the same rationale does not hold for non-ACOs. Their cognatehood becomes relevant since, in the absence of a modifier, an utterance like (27a) becomes pragmatically infelicitous, as it violates the Gricean Maxim of Quantity.

- (27) a) #*Tristram died a death.*  
 b) *Tristram died a horrible death.*

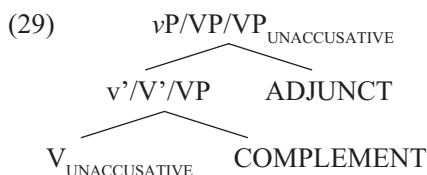
Just like manner modifiers, non-ACOs predicate the event argument directly. In the absence of the modifier, however, this result is uninformative, since the primary predicate (verb) and the secondary predicate (nominal) are identical (28a).<sup>7</sup> The relevant part of the sentence that tells us how the event of dying is happening – be it “naturally” or “horribly” – rests on the modifier, not the nominal predicate (28b).

- (28) a) #*Tristram je umro smrću.* [SC]  
 a’)  $\exists e$  [dying(e, T) & death (e)]  
 b) *Tristram je umro užasnom smrću.*  
 b’)  $\exists e$  [dying(e,T) & horrible death(e)]

7 My use of the notion «cognatehood» should be taken as a very crude shorthand. Lack of space prevents me from further discussing this, but rather than morphological or etymological relatedness, «identical» here should be understood roughly as «something that is presupposed» by the primary predicate. This is the reason why #*Opomenuo je glasom*/He warned her in a voice is infelicitous, whereas *Opomenuo je tihim glasom*/He warned her in a soft voice is fine.

## 4.5 Scarcity of Syntactic Space Revisited

Being modifiers, non-ACOs are expected to be adjuncts. This is a desirable outcome, as they cross-linguistically occur with unaccusatives.



It is completely irrelevant for the purposes of our discussion here if we think of an adjunct as a sister to V', under the direct dominance of VP and the linearisation of which is different from that of arguments or if we treat them as creatures that involve a different plane altogether. Importantly, syntactically, the position that non-ACOs occupy is never that of internal arguments.<sup>8</sup>

## 5. ZOOMING IN ON ACOS

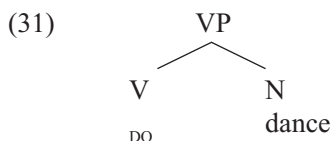
### 5.1 True Complements

Quite like regular complements, ACOs occur with strong determiners, like (30) for instance, and pass other tests other run-of-the-mill arguments do as well (recall (8) – (11)).

- (30) *Sanjal je vse sanje (ki so kdaj obstajale).* [SLO]  
 dream.prt.m.3sg AUX.3sg every dream.acc (that every existed)  
 'He dreamt every dream (that every existed)'

### 5.2 The Syntax of ACOs

So far, we have established that ACOs behave like regular complements. The question that arises now is: What is their syntactic structure? Since the dominant analysis of unergatives is the one developed by Hale and Keyser (Hale and Keyser 1993 *et seq.*), according to which the relevant unergatives involve a process of incorporation (31), this question is far from trivial.



<sup>8</sup> (29), even as a simplification, is not the only possible structure, but for the purposes of our discussion, it is sufficient.

Namely, if their analysis is correct, then there is no syntactic space for the ACO as the bare noun already occupies the position of the object. Hale and Keyser (1997) offer to solve this problem by using “index deletion”, which allows the trace left by the incorporation of *dance* into the semantically empty verb DO to be obliterated (invisible to syntax and interpretation). The process of the index deletion then allows the cognate object to be generated in the base position of the bare noun. For general theoretical reasons, this is not an optimal solution (see also Hale and Keyser 2002).

Hale and Keyser (2002) modify the account of Hale and Keyser (1997) by splitting the group of cognate objects into those that are “zero-related” and those that do not share a root with the relevant verb, then arguing that only the former, but crucially not the latter, are the results of their incorporation mechanism. Under this account, then, one could expect “dance a dance” and “dance a Mazurka” to have different syntactic structures and “derivational history”. Only in the latter case could the verb “dance” be directly generated as a “fully-fledged” verb. Something like “dance a dance” should behave like an incorporation, whereas “dance a Mazurka” should behave like a regular transitive verb. However, if put through the standard tests (as those illustrated in (8) – (11)), the two behave completely on a par. Furthermore, this theory provides no explanation as to why the verb “dance” would not be able to be directly generated in the position of the verb in the “zero-related” cases to begin with.

Moreover, one might say that the behaviour of the proverbial “die a gruesome death” and “smile a silly smile” actually argues against incorporation analysis for COs even more strongly than other pieces of data do.

Discourse referentiality (32) is typically taken to be a diagnostic of N(oun) I(ncorporation) as a syntactic phenomenon:

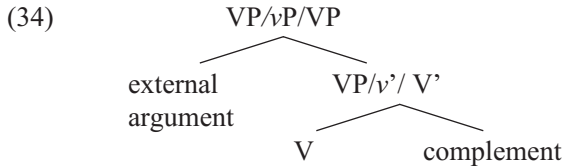
- (32) *ngii-moonahapnii mii dash ngii-giziibiiginigan* [Ojibwe]  
*n- gii- moonah -apnii -e mii dash*  
 1- PST- dig -potato -VAI and then  
*n- gii- giziibiiginig -an*  
 1 PST- wash -3pl  
 ‘I dug up potatoes, and then I washed them.’  
 (BJ, 2008-12-17, cited in Barrie & Matheu 2015:4)

However, if discourse referentiality is characteristic of NI, then non-ACOs clearly do not seem to behave like NIs, as they are never referential (recall (9)) and consequently cannot be picked up in subsequent discourse (33):

- (33) a. \**He smiled a happy smile and then her son smiled it too.*

An alternative to incorporation analyses is a transitive structure such as (34) (see also Pereltsvaig 1999).<sup>9</sup>

9 In this respect, the difference between Pereltsvaig 1999 and my account is that I argue that there is a subset of verbs that take ACOs but which are primitively unergative (see Marelj 2015 for discussion).



### 5.3 The Relevance of the Semantic Status of ACOs?

The reason we see the restriction in the examples in (36), but not in the examples in (35), is that the two types of nominal phrases are completely different semantically. Whereas non-ACOs are discussed in the previous section, 4.3 focuses on ACOs.

(35) *Pleše ples.* [ACO]  
 dance.3sg dance  
 ‘S(he) dances a dance.’

(36) *#Smeje se smehom.* [non-ACO]  
 laugh.3sg CL laugh  
 ‘#He laughed a laugh.’

Under a neo-Davidsonian view, thematic roles describe the way a participant takes part in an event. Unlike predicates, arguments then relate to an event indirectly – via thematic roles (37):<sup>10</sup>

(37) a) *Luka priča pravilno.*  
 Luka talk.3sg correctly  
 c)  $\exists e$  [talking (e) & Agent (e, Luka) & correctly(e)]

That the relation is mediated through a thematic role renders utterances in which ACOs appear always informative enough, regardless of whether ACOs are cognate or not.<sup>11</sup> Hence, no modification of these arguments is required.

Indeed, verbs that allow both ACOs and non-ACOs require modification only in the cases of non-ACOs:

(38) *Luka živi život /životom fudbalera /#živi životom.*  
 Luka live.3sg life-acc /life.instr footballer.gen /# live. life.instr

<sup>10</sup> For ease of exposition, temporal information is absent from semantic formulae throughout.

<sup>11</sup> “Informative enough” must here be understood against the background of the Neo-Davidsonian (specifically Parsons 1990) understanding of what the verb meaning is. Verbs denote one-place predicates of events and thematic roles are functions from events divided into individuals/participants of events. The ‘labels’ like Agent, Theme, Sentient inform us about the kind of involvement/the nature of the participation of an individual in the event under consideration.

Last but not least, in accordance with the hypothesis that ACOs and non-ACOs are different, note that they can also co-occur as in (39) from SC:<sup>12</sup>

- (39) (*Posle dijagnoze raka*)....  
*odlučila je da [živi život] [životom filmske zvezde]*  
 decided.prt.f.3sg AUX.3sg that live.3sg life.acc life.instr film star.gen

‘(After the cancer diagnosis)...  
 she decided to live a life of a movie star.’

#### 5.4 Making Sense of Diametrically Opposed Judgments

As evident from (38) and (39), some verbs can take both ACOs and non-ACOs. Within this analysis, one expects that ACOs of such verbs behave like regular arguments, while their non-ACOs are expected to behave like predicates. These predictions are borne out. As illustrated in (40a), for instance, while an ACO can appear with a strong determiner, a non-ACO (40b) cannot:

- (40) a) *Sanjao je svaki san (koji je ikada postojao)*. [Argument]  
 dream.prt.m.sg AUX every dream.acc (that ever existed)  
 ‘He dreamt every dream that ever existed.’
- b) \**Sanjao je svakim snom (koji je ikada postojao)*. [Predicate]  
 dream.prt.m.sg AUX every dream.instr (that ever existed)

The existence of data like (38) – (40) leads me to propose that quite like SC, English also has verbs that allow either ACOs or non-ACOs and that different judgements reflect these different options. While English lacks the initial “morphological clues” that one might be dealing with in two different types of nominals, under the relevant tests like e.g. passivisation, the data in English start behaving quite like the morphologically “robust” languages.<sup>13</sup>

12 As emphasised by R1, SLO behaves differently than SC here, as the counterpart of SC (39) is ungrammatical. What is interesting to note here is that it is not only the case that SLO disallows the co-occurrence of relevant ACOs and non-ACOs with verbs like *dream* or *live*, but also that it does not allow non-ACOs to appear with these verbs at all. It is not obvious why this is the case, but note that SLO has only a prepositional instrumental and its use is somewhat more restricted than in related languages. This seems to be relevant since the counterpart in Russian, for instance, is perfectly fine.

(i) *Ona prožila žisn’ žisn’ju spuerzvezdy*  
 She live.prf.dur life.acc life.instr superstar.gen

13 The poverty of the morphological case system in English is arguably also responsible for the inability of the two types of COs to co-occur together, giving rise to, for instance, a counterpart of (39) in English.

## 6. CONCLUSION

In this paper, I dealt with the phenomenon of cognate objects, viewed particularly through the prism of Serbo-Croatian and Slovene. The upshot of this analysis is that there are two types of COs: cognate arguments (ACOs) and cognate predicates (non-ACOs). The notion of “cognatehood” becomes relevant in the case of non-ACOs; their modification is required if utterances in which they appear are to be pragmatically felicitous. As far as ACOs are concerned, because the relation between an ACO nominal and the event variable is always mediated through a thematic role, utterances in which they occur are always informative enough. No modification of ACOs is required and their “cognatehood” is inconsequential. The status of non-ACOs as adjuncts makes the syntax of unaccusatives with which they are licensed unproblematic. As for ACOs, I argue that some of the verbs with which they occur are primitively transitive, while others are primitively unergative. Crucially, there does not seem to be good reason to treat either ACOs or non-ACOs as incorporations. An examination of morphologically robust languages such as SC and SLO facilitates an understanding of some of the puzzling properties of COCs cross-linguistically and also offers a means of explaining the disagreement regarding the judgments found in languages like English.

## References

- BARRIE, Michael/Eric MATHIEU (2015) “Noun Incorporation and Phrasal Movement.” *Natural Language and Linguistic Theory* 34/1, 1–51.
- DAVIDSON, Donald 1967. “Truth and Meaning.” In: D. Davidson (ed.), *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press, 17–36.
- HALE, Ken/Samuel Jay KEYSER (1993) “On Argument Structure and Lexical Expression of Syntactic Relations.” In: K. Hale/S. J. Keyser (eds), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, Mass: MIT Press, 53–110.
- HALE, Ken/Samuel Jay KEYSER (1997) “On the Complex Nature of Simple Predicators.” In: A. Alsina/J. Bresnan/P. Sells (eds), *Complex Predicates*. Stanford: CSLI Publications, 29–66.
- HALE, Ken/Samuel Jay KEYSER (2002) *Prolegomenon to a Theory of Argument Structure*. Cambridge, Mass: MIT Press.
- HIGGINBOTHAM, Jim (1987) “Indefiniteness and Predication.” In: E. Reuland/A. ter Meulen (eds) *The Representation of (In)definiteness*. Cambridge, Mass: MIT Press, 43–70.
- JONES, Michael Allen (1988) “Cognate Objects and the Case-Filter.” *Journal of Linguistics* 24, 89–110.
- JESPERSEN, Otto (1927) *A Modern English Grammar on Historical Principles*, Volume III (Syntax). København: Munksgaard. [Reprinted by Allen & Unwin 1954.]
- MARELJ, Marijana (2015) *On Arguments and Predicates: In the Realm of COs*. Paper presented at *FDSL 11*, University of Potsdam.



- MOLTMANN, Friederike (1989) "Nominal and Clausal Event Predicates." In: *Papers from the 25<sup>th</sup> Annual Regional Meetings of the Chicago Linguistic Society*, 300–314. University of Chicago.
- PARSONS, Terence (1990) *Events in the Semantics of English: A study in Sub-Atomic Semantics. Current Studies in Linguistics Series 21*. Cambridge, Mass: MIT Press.
- PERELTSVAIG, Asya (1999) "Cognate Objects in Russian: Is the Notion "Cognate" Relevant for Syntax?" *Canadian Journal of Linguistics* 44/3, 267–291.
- RICE, Sally (1988) "Unlikely Lexical Entries." In: *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, 202–212.
- WILLIAMS, Edwin S. (1983) "Semantic vs. Syntactic Categories." *Linguistics and Philosophy* 6, 423–446.

### Summary

#### IN THE SOUTH SLAVONIC GARDEN: LANDSCAPING THE LANDSCAPE OF ARGUMENTS AND NON-ARGUMENTS

This paper deals with morphological, syntactic, semantic, and pragmatic aspects of the so-called Cognate Object Construction with particular reference to Serbo-Croatian and Slovene. The relevance of an examination of such morphologically robust languages is manifold. It facilitates an understanding of some of the puzzling properties of the construction cross-linguistically, offers a way of explaining the noted disagreement regarding judgments found in the literature on Germanic languages such as English and also presents a clear case where (contrary to the dominant view in the literature) morphology seems to deceive, rather than inform us, about syntax. Based on a barrage of tests, I argue that there are two types of cognate objects: arguments and non-arguments. Extending the treatment of modifiers within the Davidsonian tradition to the latter, I analyse them as first-order predicates. This allows me to capture their core properties, among which is the obligatory modification, something unaccounted for in the literature. The semantic parallelism between the adverbial modifiers and non-ACOs extends to the syntax as well. Treating non-ACOs as adjuncts solves the problem of the scarcity of syntactic space that arises with unaccusative verbs that license them. ACOs, on the other hand, behave syntactically and semantically like run-of-the-mill arguments and a run-of-the-mill transitive syntax can be maintained (for a majority of them) instead.

**Keywords:** argument, cognate, predicate, unaccusative, unergative

### Povzetek

#### V JUŽNOSLOVANSKEM VRTU: RISANJE POKRAJINE GLAGOLSKIH ARGUMENTOV IN NE-ARGUMENTOV

Članek obravnava morfološke, skladišne, pomenske in pragmatične vidike t.i. zgradb s tavnološkimi predmeti s posebnim poudarkom na srbohrvaškem in slovenskem

jeziku. Obravnava morfološko bogatih jezikov je pomembna iz več razlogov. Omogoča namreč razumevanje nekaterih zapletenih lastnosti teh zgradb v več jezikih, ponudi način, kako razložiti znana razhajanja pri sodbah, ki jih najdemo v literaturi o germanskih jezikih (kot je npr. angleščina), ter služi kot jasen zgled, kako je morfologija lahko bolj zavajajoča kot informativna glede skladnje (proti prevladujočemu mnenju v literaturi). V članku na osnovi številnih testov zagovarjam trditev, da obstajata dva tipa tavnoloških predmetov, argumentni in neargumentni. Z razširitvijo obravnave modifikatorjev znotraj davidsonske tradicije na neargumente, so le-ti obravnavani kot predikati prvega reda. Ta pristop omogoča razlago njihovih temeljnih značilnosti, kot je npr. obvezna prisotnost prilastka. Pomenske vzporednice med prislovnimi določili in neargumentnimi tavnološkimi predmeti je mogoče razširiti na skladnjo. Obravnava neargumentnih tavnoloških predmetov kot prislovnih določil namreč reši problem pomanjkanja skladnjskega prostora, ki se pojavi pri netožilniških glagolih. V nasprotju z neargumentnimi tavnološkimi predmeti pa se argumentni tavnološki predmeti obnašajo kot običajni argumenti z običajno skladnjo glede glagolske prehodnosti.

**Ključne besede:** glagoski argument, tavnološki predmet, predikat, netožilniški glagol, neargumentni glagoli





## LOCALIZING CONDITIONAL CLAUSES IN THE LEFT PERIPHERY: EVIDENCE FROM MULTIPLE COMPLEMENTIZER CONSTRUCTIONS IN ROMANCE

### 1. INTRODUCTION

Adopting a cartographic approach to the structure of the left-periphery along the lines of Rizzi (1997), in this article I explore the distribution of conditional clauses in multiple complementizer constructions in Old Italian and early Italo-Romance as well as in modern Ibero-Romance. I argue that this kind of adverbial clauses could occupy in early varieties more than one specifier position within the left periphery of embedded clauses; this possibility has been lost over the centuries and is no more attested in modern Italian, where conditional clauses target just one specifier position, namely the specifier of a high functional projection within the Topic field. Modern Ibero-Romance, where complementizer doubling is still attested, patterns instead with early Italo-Romance. The discussion of the data will lead to the conclusion that in multiple complementizer constructions the function of the highest complementizer is to lexicalize the subordinating head Force<sup>o</sup>, while the lower occurrences of the complementizer mark the boundary of the two Topic subfields into which the Topic layer can be split, in the spirit of Benincà and Poletto (2004);<sup>1</sup> moreover, it will be argued that whenever an embedded clause is introduced by a single complementizer, this invariably lexicalizes Force<sup>o</sup>, the highest functional head of the left-periphery, the optionality in the lexicalization being limited to the lower Topic heads.

The article is structured as follows: in section 2, I present empirical evidence from modern Italo-Romance varieties in favour of the splitting of the Topic field into two distinct subfields, the higher of which hosting preposed adverbial clauses, the lower one hosting non-clausal topicalized constituents. In section 3, I discuss the distribution of conditional clauses with respect to the phenomenon of complementizer doubling and tripling in early Italo-Romance, pointing out that the preposed adverbial clause could target more than just one left-peripheral specifier position. In section 4, I take into account some cases of complementizer iteration from modern Ibero-Romance, arguing that the function

---

\* munaro@unive.it

1 Previous and slightly different versions of this work have been presented at the 41<sup>st</sup> *Incontro di Grammatica Generativa* (Perugia, February 2015), at the 21<sup>st</sup> *Giornata di Dialettologia* (Padua, June 2015), at the workshop *Formal Approaches to Morpho-Syntactic Variation* (Vitoria-Gasteiz, June 2015), at the 8<sup>th</sup> *SinFonIJA* conference (Ljubljana, September 2015). I would like to express my thanks to the audiences of these meetings as well as to two anonymous reviewers for insightful comments and constructive criticism.

of the lower complementizers is the delimitation of the boundary of the two main Topic subfields. Section 5 concludes the paper with some summarizing remarks.

## 2. TWO TOPIC SUBFIELDS

As discussed extensively in Munaro (2010), there are sound empirical reasons to postulate that the left-peripheral Topic field should be split into at least two distinct Topic subfields, the higher of which is dedicated to host topicalized clausal adjuncts, while the lower one hosts topicalized non-clausal phrasal constituents.

This hypothesis seems to be supported by the fact that a topicalized inverted conditional clause tends to precede a topicalized non-clausal constituent of the embedded clause, as witnessed by the following examples from modern standard Italian; both in (1a) and in (2a), which are fully grammatical, the inverted conditional clause precedes the left-dislocated internal argument of the embedded predicate, while (1b) and (2b), where the linear order is reversed, are very marginal:

(1) a) *Credo che, fosse Gianni venuto alla festa, Mario, avremmo dovuto invitarlo.*  
believe that, were John come to-the party, Mario, have-cond must invite-him  
'I believe that, had John come to the party, Mario, we should have invited.'

b) *??Credo che Mario, fosse Gianni venuto alla festa, avremmo dovuto invitarlo.*  
believe that Mario, were John come to-the party, have-cond must invite-him  
'I believe that Mario, had John come to the party, we should have invited.'

(2) a)

*Credo che, avesse Gianni rifiutato la nostra proposta, con Mario, avremmo dovuto parlare.*  
believe that, had John rejected the our proposal, with Mario, have-cond must speak  
'I believe that, had John rejected our proposal, with Mario, we should have spoken.'

b)

*??Credo che con Mario, avesse Gianni rifiutato la nostra proposta, avremmo dovuto parlare.*  
believe that with Mario, had John rejected the our proposal, have-cond must speak  
'I believe that with Mario, had John rejected our proposal, we should have spoken.'

The grammaticality contrast is somewhat less evident, but still clearly perceivable, when the preposed conditional clause is introduced by the hypothetical complementizer *se*:

(3) a)

*Credo che, se si sono incontrati prima della riunione, della tua collega, ne abbiano parlato.*  
believe that, if refl-are met before of-the meeting, about your colleague cl-have spoken  
'I believe that, if they met before the meeting, about your colleague, they have spoken.'

b)

*??Credo che della tua collega, se si sono incontrati prima della riunione, ne abbiano parlato.*  
believe that about your coll. if refl-are met before of-the meeting, cl-have spoken  
'I believe that, about your colleague, if they met before the meeting, they have spoken.'

This kind of linear restriction seems to hold crosslinguistically, as also in Paduan, a dialect spoken in the North-Eastern Italian region of Veneto, a preposed inverted conditional clause must precede both a left-dislocated constituent and a *wh*-phrase, as shown by the following examples; the only grammatical sequence is (4a), where the conditional clause precedes both the indirect object *a to sorèla* and the *wh*-item *cosa*, while the alternative orders in (4b) and (4c) give rise to ungrammaticality:

(4) a)

*Fùsselo vegnùo anca Mario, a to sorèla, cosa garissimo podùo dirghe?*  
 were-cl come also Mario, to your sister, what have-cond been-able tell-her  
 ‘If Mario had come as well, to your sister, what could you have told?’

b)

??*A to sorèla, fùsselo vegnùo anca Mario, cosa garissimo podùo dirghe?*  
 to your sister, were-cl come also Mario, what have-cond been-able tell-her  
 ‘To your sister, if Mario had come as well, what could you have told?’

c)

\**A to sorèla, cosa, fùsselo vegnùo anca Mario, garissimo podùo dirghe?*  
 to your sister, what, were-cl come also Mario, have-cond been-able tell-her  
 ‘To your sister, what, if Mario had come as well, could you have told?’

On the other hand, in Paduan the adverbial clause must follow a hanging topic constituent, as witnessed by the examples in (5), where the indirect object *Mario* functioning as hanging topic is not accompanied by the preposition and is obligatorily resumed by the pronominal clitic *ghe*; the full grammaticality of (5a), where *Mario* precedes the preposed concessive clause, suggests that the latter indeed belongs to the Topic field:

(5) a)

*Mario, anca gavesseli telefonà in tempo, no garissimo podùo dirghelo.*  
 Mario, also had-cl phoned in time, not have-cond been-able tell-him-it  
 ‘Mario, even if they had phoned in time, we couldn’t have told.’

b)

??*Anca gavesseli telefonà in tempo, Mario, no garissimo podùo dirghelo.*  
 also had-cl phoned in time, Mario, not have-cond been-able tell-him-it  
 ‘Even if they had phoned in time, Mario, we couldn’t have told.’

Adopting Benincà and Poletto’s (2004) decomposition of the Topic field into a Frame and a Thematization subfield, we can identify the landing site of topicalized adverbial clauses in the specifier of the SceneSettingP belonging to the Frame subfield, while clitic left-dislocated constituents would target the specifier of the lower LeftDisIP belonging to the Thematization subfield, as represented in (6):

(6) [ForceP [Force°][TopicP-SceneSettP conditional clause [SS°][TopicP-LeftDisIP topicalized constituent [LD°] ... ]]]

As far as embedded clauses are concerned, in modern Italian a preposed conditional clause associated to the embedded clause follows the subordinating complementizer *che*, as witnessed by the contrast between the full grammaticality of (7a) and the marginality of (7b):

(7) a)

*Credo che, se il tuo collega non verrà alla riunione, dovremo parlare con Gianni.*  
 believe that, if the your coll. not will-come to-the meeting, must-fut speak with John  
 ‘I believe that, if your colleague will not join the meeting, we will have to speak with John.’

b)

??*Credo, se il tuo collega non verrà alla riunione, che dovremo parlare con Gianni.*  
 believe, if the your coll. not will-come to-the meeting, that must-fut speak with John  
 ‘I believe, if your colleague will not join the meeting, that we will have to speak with John.’

This same ordering was attested in Old Italian, where the conditional clause usually followed the subordinating *che*:

(8) *...pensando che se sarà compagno di Dio nelle passioni,*  
 ...thinking that if will-be companion of God in-the passions,  
*Ø sarà suo compagno nelle consolazioni.*  
 will-be his companion in-the consolations  
 ‘...thinking that if he will be God’s companion in the sufferings, he will be his companion in the consolations.’

Bono Giamboni, *Libro*, chap.7, par.12

We can schematically represent the structure of (7a) and (8) as in (9), where the subordinating complementizer lexicalizes the head Force<sup>o</sup>, while the preposed adverbial clause targets the specifier of a Topic projection whose head is phonetically empty:

(9) Main clause [ForceP [Force<sup>o</sup> *che*] [TopicP conditional clause [Topic<sup>o</sup> Ø] ... ]]

However, as we will see in the next section, this was not the only option in Old Italian.

### 3. ON CONDITIONAL CLAUSES AND (MULTIPLE) COMPLEMENTIZERS IN EARLY ITALO-ROMANCE

#### 3.1 Complementizer Doubling

Beside the ordinary structure with one complementizer introducing the embedded clause, in Old Italian, by which I essentially mean 13<sup>th</sup> century Florentine, we can find numerous examples where the protasis appears sandwiched between two instances of the complementizer *che*, one preceding and one following the preposed clausal adverbial associated to the embedded clause:

(10) a)

...pe rò vi priegho in lealtade e fede che, se ttue vuoi del mio avere, che ttu ne tolghi.  
...but you-pray in loyalty and faith that, if you want of my have, that you cl-take  
'...therefore I ask you in loyalty and faith that, if you want my belongings, that you take some'  
*Libro della distruzione di Troia*, p.155, ll. 26-27

b)

...dirai (...) che, se tuo padre fu loro aspro, che tu sarai loro umile e benigno...  
...will-say that if your father was them severe, that you will-be them humble and benign...  
'...you will say that, if your father was severe to them, that you will be humble and benign to them...'  
*Novellino*, 6, ll. 37-38

c)

...e di quella cotale quantità siate sichuri da lloro, sì che se'l fatto si potesse fare,  
...and of that such amount be sure from them, so that if the fact cl-could do,  
che nnoi possiamo pagare i denari sicuramente...  
that we can pay the money surely...  
'...and of such amount be ensured by them, so that if the thing could be done, that we can surely pay the money...'  
*Lettera di Consiglio de' Cerchi II*, p. 603, ll. 24-27

d)

...ti priego che, se egli avviene ch' io muoja, che le mie cose ed ella  
...you-pray that, if it happens that I die, that the my things and she  
ti sieno raccomandate.  
you-be recommended  
'...I ask you that, if it happens that I die, that my things and she be entrusted to you.'  
*Decameron*, 2,7,84

Paoli (2007), discussing some cases of complementizer doubling in early Romance, takes the second occurrence of *che* to head the TopicP projection, which hosts the topicalized adverbial clause in its specifier:

(11) Main clause [ForceP [Force° *che1*] [TopicP conditional clause [Topic° *che2*] ... ]]

In her view, the overt realization of the complementizer in Topic° is taken to reflect a spec-head agreement relation between Topic° and the clausal constituent occupying Spec,TopicP. In the same spirit, Ledgeway (2005) – discussing the following examples of complementizer doubling from Southern Italian varieties of the 14<sup>th</sup>-15<sup>th</sup> century – interprets the first occurrence of *che* as the lexicalization of Force° and the second one as the phonetically realized trace left in the intermediate landing site Topic° by the complementizer raising from Fin° up to Force°:<sup>2</sup>

2 Also other types of adverbial clauses could appear between *che1* and *che2*, like in the following examples:  
(i) a) ...e amava sòfforte mente che a llui si era tutta via viso che quando persona neuna la sguardasse,  
...and loved so strongly that to him so was anyway shown that when person no her-watched  
che inmantenente iglile togliesse.  
that soon he-to-her took-off  
'...and he loved so intensely that to him it was shown that, when nobody was watching, that immediately he would take them off her'. *Il Tristano Riccardiano*, cap. 75, pg. 149, 25-28, from Paoli (2007)



(12) a)

*...le aveva ditto che se sua maistà voleva lo stato suo che se llo venesse*  
...him had told that if his majesty wanted the state his that refl-it- came  
*a ppigliare co la spata in mano.*

to take with the sword in hand

‘...he had told him that, if his majesty wanted his state, that he should come and take it with his sword in hand’. *Cronica* 148 v.1-2, from Ledgeway (2005)

b)

*omni raxun dichi ki si homu ad homu fa fallu, ki sia tinutu*  
all reason says that if man to man makes wrong, that be kept  
*a la debita satisfaccioni.*

to-the necessary satisfaction

‘all reason states that, if one man wrongs another, that he should remain in his debt.’

*Sposizione del Vangelo della Passione secondo Matteo* 44.14-5, from Ledgeway (2005)

As Ledgeway points out, particularly telling is the following example where the lower complementizer is followed by a focalized constituent, which reinforces the hypothesis that it lexicalizes a Topic<sup>o</sup> head, under the assumption that the landing site of focalized phrases follows all the topic-related projections (cf. Benincà and Poletto 2004):

(13)

*Eo penso bene che, se per lo tempo passato avessimo voluto monstrare lo nostro ardire,*  
I think well that, if for the time passed had wanted to show the our bravery  
*che DERITAMENTE avessimo mandato ad asseyare Troya.*

that straightaway had sent to siege Troy

‘I indeed think that, if in the past we had wanted to show our bravery, that STRAIGHTAWAY we would have gone to siege Troy.’

*Libro della Destructione di Troya* 140.21-3, from Ledgeway (2005)

From the data reported in this section, we can conclude that in early Italo-Romance varieties *if*-clauses – and adverbial clauses in general – were among the most plausible

---

b) *volimo et commandamote che, mantinente che per lictere nostre senteray essere nuy o a Melfe*  
want and order-you that, as soon as for letters ours will-hear to-be we or at Melfi  
*o a Troya, che dige ad nuy senza dimoranza personalmente venire.*

or at Troy, that should to us without lingering personally come

‘we want and order you that, as soon as you hear thorough one of our letters that we are in Melfi or Troy, that you should personally come at once to us.’ *Lettera del re Luigi d’Angiò-Taranto...* 12-4, from Ledgeway (2005)

c) *serrà bisogno che, dove ilo non mecte exemplo per lo quale poza bene essere intiso lo suo dicto,*  
will-be need that, where he not puts example for the which can well be understood the his word  
*che eo mecta exemplo e declaracione per manifestare lo intendimento suo.*

that I put example and declaration for show the intention his

‘it will be necessary that, wherever he fails to provide as example by which his words can be clearly understood, that I give an example and declaration in order to make his intention clearly understood.’

*Libro de lu Dialagu de Sanctu Gregoriu* 3.1-3, from Ledgeway (2005)

candidates to fill the position sandwiched between the two occurrences of the complementizer, the higher one heading Force<sup>o</sup>, the lower one heading a Topic projection.<sup>3</sup>

### 3.2 A higher position for conditional clauses

Another possibility attested in Old Italian, and not attested in modern Italian, is the presence of a single phonetically realized complementizer after the conditional clause associated to the embedded clause, like in the following examples:

(14) a)

*...e dico ben Ø, se l voler non mi muta, ch'eo stesso li uccidrò, que' sconoscenti.*  
 ...and say well, if the will not me-changes, that I myself them-will-kill, those louts  
 '...and I say well, if I do not change my will, that I will kill them myself, those louts'

Dante, *Rime*, 8, vv. 13-14

b)

*...per ch'io son certo Ø, se ben la difendo nel dir com'io la 'ntendo,*  
 ...for that I am certain, if well her-defend in saying how I her-mean,  
*ch' Amor di sé mi farà grazia ancora.*  
 that Love of himself me-will-do grace again  
 '...therefore I am certain, if I defend her well in saying how I mean her, that Love will favour me again.'

Dante, *Rime*, 30, vv. 17-19

c)

*Ma so bene Ø, se Carlo fosse morto, che voi ci trovereste ancor cagione...*  
 but know well, if Charles were dead, that you cl-find-cond still reason...  
 'But I know well, if Charles had died, that you would find new reasons...'

Rustico Filippi, *Sonetti*, 3, vv. 9-10

3 In the few modern Italo-Romance varieties still displaying complementizer doubling (cf. Paoli 2007), conditional clauses cannot occur any more between the two complementizers. Under the present approach, the impossibility in modern Italo-Romance to sandwich a clausal adjunct between two complementizers can be interpreted as a consequence of a diachronic process of downward reanalysis to the effect that *che2*, originally lexicalizing the higher head SceneSett<sup>o</sup>, has been reanalyzed by the speakers – due to the structural ambiguity of the linear string – as the lexicalization of the lower head LeftDisl<sup>o</sup>, namely the one associated with the specifier position hosting left-dislocated phrasal constituents, as represented in (i):

- (i) a) [ForceP [Force<sup>o</sup> *che1*] [SceneSettP adverbial clause [SS<sup>o</sup> *che2*] [LeftDislP [LD<sup>o</sup>]]]  
 b) [ForceP [Force<sup>o</sup> *che1*] [SceneSettP adverbial clause [SS<sup>o</sup>] [LeftDislP [LD<sup>o</sup> *che2*]]]  
 c) [ForceP [Force<sup>o</sup> *che1*] [SceneSettP [SS<sup>o</sup>] [LeftDislP topicalized phrase [LD<sup>o</sup> *che2*]]]

Interestingly, as pointed out by an anonymous reviewer, a Slovenian dialect that has language contact with Italian still displays conditional clauses in complementizer doubling structures, like in the following example taken from Plesničar (2015):

- (ii) *Koga si rekel, da če kaj ni jasno, da naj vprašajo?*  
 who aux say that if something not clear that should ask  
 'Who did you say that they should ask if something isn't clear?'

According to Meszler and Samu (2010), in all these examples the conditional clause preceding *che* clearly refers to the content of the subordinate clause, and not to the matrix clause, despite appearance, so that we are obliged to assume that the clausal adjunct preceding *che* belongs indeed to the left periphery of the embedded clause; in their view, the phonetic realization of the lower complementizer makes the overt realization of the higher one superfluous. I would like to propose a different analysis for the examples in (14), suggesting that the complementizer *che* is in fact the lexicalization of Force<sup>°</sup> and that the conditional clause has raised to a higher structural position, namely the specifier of ForceP, crossing over the complementizer:<sup>4</sup>

(15) Main clause [ForceP conditional clause [Force<sup>°</sup> *che*] [TopicP [Topic<sup>°</sup> Ø] ... ]]

This alternative structural analysis allows us to formulate a generalization, namely that whenever we find an embedded clause introduced by a single complementizer, this invariably lexicalizes Force<sup>°</sup>, the highest functional head of the left-periphery, and the optionality in the lexicalization concerns exclusively the lower Topic heads.

### 3.3 Complementizer tripling in early Italo-Romance

In early Italo-Romance a further possibility was sporadically attested, namely the presence of three complementizers, with topicalized lexical material appearing between them. Vincent (2006) reports the following example, where a conditional clause occurs between the highest and the intermediate *che*, while a heavy topicalized constituent, the subject of the embedded clause, appears between the intermediate and the lowest *che*:

(16)

*Ancora statuemo e ordenamo che se alcuna persona de la dita Confraria fosse maroto*  
 still establish and order that if any person of-the said Company were dead  
*che subitamente quello chi saverà de quello ditto fraello maroto che ello lo debia denuntiar*  
 that soon that who will-know of that said brother dead that he it- must tell  
*a lo prior*  
 to-the prior

‘We further establish and order that, if any person of the mentioned Company died, that soon who will know about that dead brother, that he should announce it to the prior’

*Statuti della Compagnia dei Caravana del porto di Genova, 1340*

Capitalizing on the recursive nature of Topic projections postulated by Rizzi (1997), we could analyze this example as follows, namely with the second and the third instance of the complementizer lexicalizing the heads of recursive Topic projections,

4 I will remain agnostic here as to the trigger for the movement of the conditional clause to Spec,ForceP, simply suggesting that this might be a syntactic device to bring the adverbial clause in a local relation of spec-head agreement with the complementizer *che* sitting in Force<sup>°</sup>.

where the conditional clause occupies a specifier position higher than the one targeted by the topicalized constituent:

- (17) Main clause [ForceP [Force° *che1*] [TopicP conditional clause [Topic° *che2*] [TopicP topicalized constituent [Topic° *che3*] ... ]]]

On the other hand, Ledgeway (2005) reports the following example of complementizer tripling where the relative order of conditional clause and heavy topicalized constituent is reversed:

- (18)  
*Pregove, madama, per l'amor di Dio, che de chilli dinare che eo agio vostri*  
 pray-you madam for the love of God that of that money that I have yours  
*che si non vi fusse troppo sconço che mi 'ndi impristiti una unça.*  
 that if not you-were too-much inconvenience that me-cl- lend an 'uncia'  
 'I ask of you, my lady, for the love of God, that, from that money of yours that I hold,  
 that, if you were not to find it too inconvenient, that you should lend me an 'uncia' of it,'  
*Lettera del tesoriere Tommasino da Nizza a Lapa Acciaiuoli, 1353*

As one can clearly see, in this case the conditional clause appears between the intermediate and the lowest *che* and follows the topicalized prepositional phrase, therefore it arguably occupies the specifier of a lower Topic projection, as represented in (19):

- (19) Main clause [ForceP [Force° *che1*] [TopicP topicalized constituent [Topic° *che2*] [TopicP conditional clause [Topic° *che3*] ... ]]]

If this analysis is on the right track, it provides evidence for a possible location of the conditional clause in a relatively low Topic position within the left periphery of the embedded clause in early Italo-Romance.

Summing up, in this section we have seen that in early Italo-Romance conditional clauses could occupy up to three different specifier positions within the left-periphery of the embedded clause, namely a low Topic position, like in (18), a higher Topic position, like in (16) and (10), and the specifier of ForceP, like in (14).

#### 4. ON COMPLEMENTIZER ITERATION IN MODERN IBERO-ROMANCE

Within the Romance domain, the possibility for a topicalized constituent or an *if*-clause to intervene between two instances of *que* is robustly attested in modern Ibero-Romance, as witnessed by the following examples:

- (20) a) *Acho que se lhe ligasses que tudo se resolveria.*  
 think that if him/her called that all itself-solve  
 'I think that if you called him/her everything would turn out fine.'

- b) *Espero que a Ana que traga o livro.*  
 hope that the Ana that brings the book  
 ‘I hope that Ana brings the book.’ European Portuguese (from Mascarenhas 2007)
- (21) a) *Me dijeron que si llueve, que viene Guillermo.*  
 me told that if rains that comes William  
 ‘They told me that if it rains, William will come.’  
 b) *Susi dice que, a los alumnos, que les van a dar regalos.*  
 Susi says that to the students that cl- go to give presents  
 ‘Susy says that they are going to give the students presents.’  
 Spanish (from Villa-Garcia 2012)
- (22) a) *La secretària em va dir que si pagava l’ import abans d’una setmana,*  
 the secretary me told that if paid the amount before a week  
*que encara em podia matricular.*  
 that still me could register  
 ‘The secretary told me that if I paid the amount before a week I could still register.’  
 b) *Ha dit que els convidats, que estan asseguts a taula.*  
 has said that the guests that stay seated at table  
 ‘(S)he has said that the guests are seated at the table.’  
 Catalan (from González i Planas 2013)

According to Mascarenhas (2007), Villa-Garcia (2012), and González i Planas (2013), in the complementizer doubling structures in (20)-(22) *que1* lexicalizes the Force<sup>o</sup> head, while *que2* lexicalizes a Topic<sup>o</sup> head which hosts the topicalized adverbial clause or the topicalized phrase in its specifier, as schematically represented in (23):<sup>5</sup>

- (23) Main clause [ForceP [Force<sup>o</sup> *que1*] [TopicP adverbial clause/topicalized constituent [Topic<sup>o</sup> *que2*] ... ]]

Villa-Garcia (2012) also discusses the following example, where the conditional clause follows *que2*, and a topicalized argumental prepositional phrase appears sandwiched between *que1* and *que2*:

- (24) *Me dijeron que a la fiesta, que si llueve, no van a ir.*  
 me said that to the party that if rains not go to go  
 ‘They told me that they are not going to the party if it rains’

5 If we accept the correctness of this approach, then the fact that preposed protases can easily enter the complementizer doubling construction can be seen as an empirical argument in favour of the hypothesis that they belong to the Topic field of the associated (embedded) clause, as independently proposed by Munaro (2005). For a recent exhaustive analysis of the distributional properties of central and peripheral adverbial clauses the reader is referred to Haegeman (2012).

According to Villa-García, the structural representation of (24) is the following, where the conditional clause occupies the specifier of the lower Topic projection (a possibility that was attested in Old Italian), whose head remains phonetically unrealized:

- (25) Main clause [ForceP [Force° *que1*] [TopicP *a la fiesta* [Topic° *que2*] [TopicP *si llueve* [Top° Ø] ... ]]]

Both in Spanish and in European Portuguese are also attested cases of complementizer tripling, like the following, where between each pair of complementizers appear topicalized constituents interpretively linked to the embedded clause:<sup>6</sup>

- (26) a) *Dijo que el dinero, que a Juan, que se lo mandaban por correo.*  
 said that the money that to Juan that cl-cl sent for mail  
 ‘He said that they were sending John the money through the mail.’  
 Spanish (from Escribano 1991)
- b) *Acho que amanhã que a Ana que vai conseguir acabar o trabalho.*  
 think that tomorrow that the Ana that will manage to finish the assignment.  
 ‘I think that tomorrow Ana will manage to finish the assignment.’  
 European Portuguese (from Mascarenhas 2007)

For an example like (26b) Mascarenhas (2007) proposes that the second and the third complementizer lexicalize the heads of recursive Topic projections:

- (27) *Acho* [ForceP [Force° *que*] [TopicP *amanhã* [Topic° *que*] [TopicP *a Ana* [Topic° *que*] ... ]]]

Interestingly, as pointed out by Mascarenhas (2007), in European Portuguese the possibility of having four complementizers co-occurring within the same clause with three topicalized constituents appearing in the embedded left periphery is excluded, as witnessed by (28a), and slightly marginal is also the appearance of a single complementizer preceding three topicalized constituents, like in (28b):

- (28) a) *??Duvido que ontem que o Pedro que à Ana que lhe tenha telefonado.*  
 doubt that yesterday that the Pedro that to the Ana that cl.dat have called  
 ‘I doubt that yesterday Pedro called Ana.’

6 The template exemplified in (26), namely the possibility of having multiple complementizers with topicalized material intervening in between each pair, is reported to be ungrammatical by Demonte and Fernández-Soriano (2009), who provide the following example:

(i) *Te pido que a tu padre (\*que) en este momento (\*que) esa mentira (que) no se la digas.*  
 you ask that to your father (that) at this moment (that) that lie (that) not cl-cl tell  
 ‘I ask you not to tell that lie to your father at this moment.’

This might suggest that in this case the three topicalized constituents are all contained within the same Topic subfield, arguably the lower one.

- b) ?*Duvido que ontem o Pedro à Ana lhe tenha telefonado.*  
 doubt that yesterday the Pedro to the Ana cl.dat have called  
 ‘I doubt that yesterday Pedro called Ana.’

The deviance of (28a), as well as the fact that we do not find in early Italo-Romance any examples containing four (or more) instances of the complementizer *che*, suggests that crosslinguistically the maximal number of occurrences of the complementizers within the same clause is three.<sup>7</sup> This can be interpreted as indicating that, excluding the highest complementizer, which uncontroversially lexicalizes Force<sup>o</sup>, the function of the two following complementizers is to mark the lower boundary of the two main Topic subfields identified above, namely the higher one, endowed with scene setting properties, and the lower one, devoted to the thematization of argumental constituents.

## 5. CONCLUSION

Based on evidence from the distributional properties of preposed conditional clauses in multiple complementizer structures in Romance, in this article I have tried to argue for a splitting of the Topic field into two subfields, along the lines of Benincà and Poletto (2004); in modern Italian the higher Frame subfield, endowed with a scene setting function, is the landing site of preposed adverbial clauses or of time adverbials; the lower subfield, devoted to the thematization function, hosts primarily topicalized non-clausal argumental constituents of the embedded clause. The respective lower boundary of these two subfields could be lexicalized in early Italo-Romance by the second (and third) occurrence of the complementizer in multiple complementizer constructions, a possibility which is still attested in modern Ibero-Romance languages. From a diachronic perspective, it is possible to determine a relation between the presence of complementizer doubling on the one hand and the relative freedom of placement of conditional clauses on the other (the situation of early Italo-Romance and modern Ibero-Romance); the loss of complementizer doubling in most modern Italo-Romance

7 As pointed out by Mascarenhas (2007), the situation is somewhat different with the interrogative complementizer *se*, which can easily appear twice in the same clause, like in (i), but can less easily appear three times, as witnessed by the slight deviance of (iia), although (iia) sounds better than (iib), where two topicalized constituents appear sandwiched between the two occurrences of *se*:

(i) *Não sei se o João (se) vai chegar a horas.*

not know if the João if will arrive at hours

‘I don’t know if João will arrive on time’

(ii) a) ?*Não sei se amanhã se o Pedro se consegue entregar o trabalho.*

not know if tomorrow if the Pedro if manages hand in the assignment

‘I don’t know if tomorrow Pedro will manage to hand in the assignment.’

b) ?\**Não sei se amanhã o Pedro se consegue entregar o trabalho.*

not know if tomorrow the Pedro if manages hand in the assignment

‘I don’t know if tomorrow Pedro will manage to hand in the assignment.’

For a possible implementation of the interrogative complementizer *se* within the split left-periphery the reader is referred to Rizzi (2001).

varieties entails a gradual shift to a stricter localization of conditional clauses within the left periphery.

## References

- BENINCÀ, Paola/Cecilia POLETTI (2004) “Topic, Focus and V2 – Defining the CP sublayers.” In: L. Rizzi (ed.), *The structure of IP and CP*. Oxford/New York: Oxford University Press, 52–75.
- DEMONTE, Violeta/Olga FERNANDEZ-SORIANO (2009) “Force and finiteness in the Spanish complementizer system.” *Probus* 21/1, 23–49.
- ESCRIBANO, José Luis Rafael G. (1991) *Una teoría de la oración*. Oviedo: Servicio de Publicaciones de la Universidad de Oviedo.
- GONZALEZ I PLANAS, Francesc (2013) *On quotative recomplementation: Between pragmatics and morphosyntax*. Manuscript, University of Girona.
- HAEGEMAN, Liliane (2012) *Adverbial clauses, main clause phenomena, and the composition of the left periphery*. Oxford/New York: Oxford University Press.
- LEDGEWAY, Adam (2005) “Moving through the left periphery: The dual complementizer system in the dialects of Southern Italy.” *Transactions of the Philological Society* 103/3, 339–396.
- MASCARENHAS, Salvador (2007) *Complementizer doubling in European Portuguese*. Manuscript, Amsterdam ILLC/ NYU.
- MESZLER, Lenka/Borbála SAMU (2010) “Le strutture subordinate.” In: G. Salvi/L. Renzi (eds), *Grammatica dell’italiano antico*. Bologna: Il Mulino, 763–782.
- MUNARO, Nicola (2005) “Computational puzzles of conditional clause preposing.” In: A.M. Di Sciullo (ed.), *UG and external systems*. Amsterdam: John Benjamins, 73–94.
- MUNARO, Nicola (2010) “Towards a hierarchy of clause types.” In: P. Benincà/N. Munaro (eds), *Mapping the left periphery*. Oxford/New York: Oxford University Press, 125–162.
- PAOLI, Sandra (2007) “The fine structure of the left periphery: COMPs and subjects. Evidence from Romance.” *Lingua* 117/6, 1057–1079.
- PLESNIČAR, Vesna (2015) *Complementizer doubling around embedded clauses in Slovenian*. Abstract. <http://skrabcevi-dnevi.zrc-sazu.si/Program.aspx>
- RIZZI, Luigi (1997) “The fine structure of the left periphery.” In: L. Haegeman (ed.), *Elements of grammar*. Dordrecht: Kluwer, 281–337.
- RIZZI, Luigi (2001) “On the position ‘Int(errogative)’ in the left periphery of the clause.” In: G. Cinque/G. Salvi (eds), *Current studies in Italian syntax. Essays offered to Lorenzo Renzi*. Amsterdam: Elsevier, 287–296.
- VILLA-GARCIA, Julio (2012) “Recomplementation and locality of movement in Spanish.” *Probus* 24/2, 257–314.
- VINCENT, Nigel (2006) “Il problema del doppio complementatore nei primi volgari d’Italia.” In: A. Andreose/N. Penello (eds), *LabRomAn: Giornata di lavoro sulle varietà romanze antiche*. Padova: Università di Padova, 27–42.



## Summary

### LOCALIZING CONDITIONAL CLAUSES IN THE LEFT PERIPHERY: EVIDENCE FROM MULTIPLE COMPLEMENTIZER CONSTRUCTIONS IN ROMANCE

This article analyzes the distribution of conditional clauses in multiple complementizer constructions, showing that preposed adverbial clauses could occupy in early Italo-Romance varieties different specifier positions within the left periphery of embedded clauses, a possibility that has been lost with time in Italo-Romance but is still attested in modern Ibero-Romance. It is further argued that in multiple complementizer structures the highest complementizer invariably lexicalizes the head Force<sup>o</sup>, while the other occurrences of the complementizer mark the lower boundary of the main two Topic subfields.

**Keywords:** complementizer doubling, conditional clause, Ibero-Romance, Italo-Romance, Topic field

## Povzetek

### SKLADENJSKI POLOŽAJ IN RAZVRSTITEV POGOJNIH ODVISNIKOV V LEVI PERIFERIJ Z VIDIKA ROMANSKIH VEČVEZNIŠKIH STRUKTUR

Članek obravnava distribucijo pogojnih odvisnikov v večvezniških strukturah in pokaže, da so se v zgodnjih italo-romanskih jezikih predpostavljeni prislovni odvisniki pojavljali na različnih položajih znotraj leve periferije vloženi odvisnikov. Ta skladijski pojav se v sodobnih italo-romanskih jezikih ni ohranil, vendar ga še vedno lahko opazimo v iberoromanski jezikovni skupini. Skladijska razčlemba večvezniških struktur pokaže, da najvišje ležeči veznik vedno leksikalizira jedro funkcijske zveze Force, medtem ko nižje ležeči vezniki označujejo mejo med glavnima nižje ležečima podpoljema funkcijske zveze Topik.

**Ključne besede:** podvajanje veznika, pogojni odvisnik, iberoromanski jeziki, italo-romanski jeziki, Topikalno polje



## SIGN ORDER IN SLOVENIAN SIGN LANGUAGE LOCATIVE CONSTRUCTIONS<sup>\*\*</sup>

### 1. INTRODUCTION

Cross-linguistically, locative constructions look very much alike, because their word order seems to be dependent on how certain semantic features, namely the size, mobility, and animateness, are valued on participants. Bigger, immobile, and usually non-animate participants (grounds) that perform spatial anchoring often tend to precede smaller, mobile, and usually animate participants (figures). Therefore, in linguistics, the terms “figure” and “ground” are used to describe the system by which language establishes one participant as a reference point for another participant.

In oral languages, the combination of a spatial adposition with an existential verb is habitually used to encode the relation between the figure and the ground. In sign languages, the existential verb is often not present and spatial information is not necessarily vehiculated by overt spatial adpositions. Instead, the location of the figure with respect to the ground is encoded within a morphologically complex verb sign, which is usually represented by a classifier predicate.

Classifiers in sign languages are meaningful hand configurations that denote a salient characteristic of their referent. They are not independent signs unless combined with a movement subcomponent in order to form classifier predicates. The movement subcomponent of such predicate is said to represent its root, while the handshape (a classifier) is a bound morpheme that may refer back to the participant(s) in the described event. In various sign languages, it has been observed that classifier predicates may influence the constituent order of the transitive sentence. In languages with a basic SVO word order, such as Jordanian (Hendriks 2007), Colombian

---

\* matic.pavlic@ung.si

\*\* This study is based on subsections 7.2 and 7.3 of my doctoral thesis. It was partially supported by the Scholarship for Slovenian Citizens for Postgraduate Study Abroad, number 11010-716/2012, COST Action IS1006 Summer School Grant IS1006-060614-044777, and Ca' Foscari internal funds for PhD mobility. The Slovenian Association of Deaf Clubs (Zveza Društev Gluhih in Naglušnih Slovenije, <http://www.zveza-gns.si/>) helped me to establish and maintain contact with SZJ interpreters (Veronika Ciglar, Natalija Rot, and Natalija Spark) and native deaf signers. For their generous collaboration, I am greatly indebted to my deaf informants Maja Kuzma, Nina and Jan Orešnik, Robert Veršič, and Valerija Škof. I am also grateful to my supervisor Chiara Branchini, my Amsterdam host Roland Pfau, anonymous reviewers, and audience members of FEAST 3 in Venice, SinFonIJA 7 in Graz, SinFonIJA 8 in Ljubljana, and Obdobja 34 in Ljubljana for their kind suggestions and useful comments. All remaining errors are my own.

(Oviedo 2003), Russian (Kimmelman 2012), and Hong Kong Sign Language, they yield an SOV word order, for example.

In locative constructions, the movement subcomponent of the classifier predicate is modified according to the locations at which the ground and figure are articulated in signing space so that the predicate movement starts in the location where the figure is produced and ends in the location where the ground is produced. In this way, a complex meaning glossed as BE-LOCATED+CL(handshape) is encoded. In SOV sign languages such as Irish (Johnston et al. 2007), Italian (Laudanna 1987), and Sign Language of the Netherlands (Coerts 1994), as well as in SVO sign languages such as American (Liddell 1980), Russian (Kimmelman 2012), Croatian (Milković 2007), Australian, and Flemish (Johnston et al. 2007) Sign Language, locative sentences are reported to display non-basic OSV word order.<sup>1</sup>

In this study, I focus on a locative construction that signers of Slovenian Sign Language (SZJ) use in order to encode a locative relation between two participants. I carefully describe the construction and attempt to determine whether or not its non-basic word order is a result of the effect of a classifier predicate, a locative environment, or both. In section 2, I present the methods of my research. In section 3, I analyse locative constructions in SZJ: firstly, those with classifier predicates (3.1), and then those with non-classifier predicates (3.2). In section 4, I explore locative arguments: the non-manuals accompanying the ground (4.1) and the distributivity of the Figure (4.2). Finally, in section 5, I hypothesise how the surface word order of SZJ locative constructions is derived and revisit the research questions that read as follows:

- RQ1 How does SZJ encode locative information: by using a locative adposition within an adposition phrase (as in many oral languages) or by modulating a spatially-agreeing predicate – as in many sign languages?
- RQ2 Which types of verbs may represent a predicate in SZJ locative constructions: classifier and/or non-classifier predicates?
- RQ3 What is the basic word order in SZJ locative constructions with respect to the type of predicate used?

## 2. METHODS

### 2.1 Slovenian Sign Language (SZJ)

SZJ is the language of the Deaf community in Slovenia. It has received close to no attention by linguists and is also virtually unknown to the majority of Slovenian population. The language is estimated to be used by 900 deaf signers as their primary means

---

1 For a broader and more detailed introduction to sign language phonology and agreement, see relevant chapters such as “Chapter 8: Classifiers and Chapter” in *Sign Languages: A Cambridge Language Survey* and “Chapter 12: Word Order”, among others, in *Sign Languages (Handbooks of Linguistics and Communication Science 37)* and the 2012 article “On the Syntax of Spatial Adpositions in Sign Languages” by Roland Pfau and Enoch O. Aboh.

of communication and by as many as 1600 signers altogether (Vintar et al. 2012). The majority of these deaf signers are at least to some extent familiar with both SZJ and spoken Slovenian.

## 2.2 Subjects

Out of six informants included in my research, three are female and three male. They are L1 deaf signers, ranging in age from 25 to 35, and come from families in which at least one parent is deaf. They are all members of local Deaf clubs and are well-integrated into the Deaf community. Two of these signers are siblings. All informants collaborated voluntarily and were not paid for their participation. They approved the publication of their data (transcriptions, clips, and stills) for research purposes.

## 2.3 Elicitations and Materials

In eliciting the data, I followed the influential work by Volterra et al. (1984), who first introduced the Picture Description Task (PDT) to sign language studies in order to investigate the word order of Italian Sign Language (LIS). Since they paid special attention to locative constructions, this methodology was especially suitable for my research. Nevertheless, I designed the stimuli (photos and illustrations of still lifes) myself. My informants were shown these stimuli one by one on a computer screen and were asked to describe the depicted situations to the interpreter/deaf co-signer.

## 3. LOCATIVE PREDICATION

In almost all SVO and SOV sign languages for which locative constructions have been studied in detail, it has been discovered that they trigger a non-basic word order as compared to the word order of arguments in transitive sentences. In (1), this is illustrated for an SOV sign language: Sign Language of the Netherlands (NGT; Coerts 1994). In order to describe the situation, the NGT signer articulates the sign TABLE before the sign BALL. Lastly, the relationship among them is set by the classifier predicate, which starts in the location where BALL is produced and ends in the location where the sign TABLE is produced. This yields the OSV word order.

- (1) TABLE<sub>A</sub> BALL<sub>B</sub> BE-LOCATED+CL<sub>A</sub> (Coerts 1994: 65; NGT)  
‘A/the ball is under a/the table’
- (2) FENCE CAT SLEEP (Liddell 1980: 91-100; ASL)  
‘A/the cat sleeps on a/the fence’
- (3) CHAIR BIG CAT IX SIT+CL (Kimmelman 2012: 37; RSL)  
‘A/the big cat sits on a/the chair’

In example (1), a complex locative meaning is encoded without using any spatial adpositions. Indeed, in the literature on sign languages, it is commonly assumed that

many sign languages lack spatial adpositions and encode locative information through spatially modified predicates instead. These predicates connect two distinct locations in space by movement and may be modified so that their movement begins in the location at which the figure has been articulated (location ‘B’ in example (1)) and ends in location at which the ground has been articulated (location ‘A’ in example (1)). In addition, such a predicate may be represented by a classifier or non-classifier predicate. In (2), for example, the American Sign Language (ASL) signer uses the citation form of the verb “sleep”. In the Russian Sign Language (RSL) example (3), on the other hand, the signer uses an agreeing verb SIT+CL articulated with a classifier handshape denoting the type of entity that is seated: namely, a four-legged animal. In fact, cross-linguistically, signers opt for the latter, so that locative predicates tend to be signed with a classifier predicate. Consequently, when researching locative constructions in sign languages, the majority of researchers focus on locative constructions employing classifier predicates, to the exclusion of locative constructions with non-classifier predicates. I believe this is a research gap because, compared to non-classifier predicates, classifier predicates *per se* may trigger non-basic word orders. In subsection 3.1, I analyse SZJ locative constructions with classifier predicates and compare them to locative constructions with non-classifier predicates in subsection 3.2.

### 3.1 Classifier Predicate

SZJ may encode the spatial location of an entity or an event in various ways. In this section, I present the possibility that was used most frequently and most consistently among my informants. I demonstrate that it is analogous to the locative constructions reported for many other unrelated sign languages, such as American (Liddell 1980), Croatian (Milković et al. 2007), Australian, Flemish, Irish (Johnston et al. 2007), Russian (Kimmelman 2012), and Sign Language of the Netherlands (Coerts 1994). The examples presented below were elicited by the stimulus: a photo that depicts a river with two rows of moored vessels aligned with its left and right bank, respectively. In both examples, the ground (noun phrase RIVER1 or RIVER2) is set first. It is non-manually marked with raised eyebrows. Then the figure (quantified noun phrase MANY VESSEL) is introduced. Finally, the relationship between them is established with a predicate BE-LOCATED+CL(B)+DM. The predicate sign is reduplicated (glossed as DM, distributivity morpheme) in order to denote the number of vessels aligned along the signing space in which the sign RIVER was produced. The word order is ground-figure-predicate.

In SZJ, the noun VESSEL is a two-handed sign, articulated by a b-configuration parallel to the lateral axis. The hands are oriented towards each other and joined at the fingers. They iconically represent the prow of a ship – as in examples (4) and (5). In (6), however, the very same hand configuration is also used for the predicate. Both signs (the noun and verb) are superficially similar with regard to their handshape, but they can easily be set apart by their movement subcomponent. The sign for the noun VESSEL is produced with two repeated circular movements in a given location in space, while the sign for the predicate BE-LOCATED+CL(B) is produced with one straight short movement ending with a hold in a given location in space. Compare the aforementioned predicate

with the predicate BE-LOCATED+CL(B) in examples (4-5). The latter is a one-handed sign produced with the b-configuration, which refers to vehicles in general, while the former was produced as a two-handed sign in a b-configuration, representing vessels in particular. Thus, a set of various different meaningful handshapes can be used in order to produce the predicate sign BE-LOCATED+CL in (4), (5), and (6). Since verbal classifier systems in both sign and spoken languages generally allow for variability in the choice of a classifier, such that more than one particular verbal classifier may combine with the same verbal root (Slobin et al. 2002 for sign languages, Aikhenvald 2000 for spoken languages), this determines SZJ predicate BE-LOCATED+CL as a classifier predicate.

(4)



\_\_\_\_\_re

<sup>A</sup>RIVER1<sub>B</sub>

MANY

VESSEL

<sup>A</sup>BE-LOCATED(B)<sub>B</sub>+DM

‘There are many vessels on the river.’

(SZJ; loc10n)

(5)



\_\_\_\_\_re

<sup>A</sup>RIVER2<sub>B</sub>

MANY

VESSEL

<sup>A</sup>BE-LOCATED(B)<sub>B</sub>+DM

‘There are many vessels on the river.’

(SZJ; m25)

(6)



\_\_\_\_\_ re

<sub>A</sub> RIVER<sub>B</sub>3

VESSEL

<sub>A</sub> BE-LOCATED(B)<sub>B</sub> +DM

‘There are (some) vessels on the river.’

(SZJ; v7)

I can now distinguish non-classifier predicates from classifier predicates by employing a simple test. A classifier predicate changes its handshape with respect to certain salient characteristics of the predicated participant, while a lexical verb is not affected in such a way. In the next subsection, I turn to SZJ locative constructions featuring non-classifier predicates.

### 3.2 Non-Classifier Predicates

In this subsection, I examine locative constructions with non-classifier predicates in SZJ. Again, the ground constituent is fronted – similar to locative constructions with a classifier predicate. The figure, on the other hand, is produced only after the predicate – contrary to locative constructions with a classifier predicate.

The SZJ verb LAY is a two-handed sign produced in l-configuration, oriented away from the signer and pointing upwards. During the movement, the arms are fixed, except for the wrists. The wrists hinge, and as the hands move downwards, the fingertips circumscribe an arc movement until they point forward. Since the verb form does not change with respect to the laying participant and since it does not denote its salient characteristics, it is not a classifier predicate, but rather a lexical verb. In example (7), this movement ends in the r-locus where TABLE was previously signed. The sentential structure is different from the one attested in SZJ locative predicates with a classifier predicate. The sentence opens with the ground TABLE, which is in turn followed by the verb LAY and finally ends with the figure POTATO. The word order is ground-predicate-figure.



(7)



\_\_\_\_\_re

TABLE

LAY

POTATO

‘A/the potato lays on a/the table.’

(SZJ; v14)

### 3.3 Interim Summary

In the literature on word order, figures are usually linked to subjects (S), while grounds are linked to objects (O). The figure-ground word order may be compared to the SO word order pattern, while the ground-figure word order may be compared to the OS word order pattern. If we apply it to the SZJ data presented above, we can conclude that the word order in SZJ locative constructions appears to be OSV for classifier predicates and OVS for non-classifier predicates. Note that to compare the correlation pair ground-figure to the correlation pair object-subject does not necessarily mean that figures and grounds indeed take subject and object roles in SZJ locative constructions. I will examine the syntactic functions of the ground and the figure constituent in the next section (4.3).

## 4. LOCATIVE ARGUMENTS

Above, I presented locative constructions with classifier predicates, and I compared them to locative constructions with non-classifier predicates. The former feature an OSV word order, while the latter feature an OVS word order. Both begin with the ground, which is non-manually marked with raised eyebrows. What do raised eyebrows mark? In subsection 4.1, I analyse the ground in SZJ as a constituent that is fronted in non-argumental movement. In subsection 4.2, I use a distributive-morpheme test to demonstrate that figures are base generated as internal arguments.

### 4.1 Ground

In all SZJ locative constructions that I have presented so far, the ground occupied the first position in the clause, regardless of the type of predicate (classifier or non-classifier predicate). In this subsection, I demonstrate that this is a pre-subject position and attempt to determine its category.



In addition to the canonical locative construction (10), I managed to record a locative predicate with a fronted ground and an agent included in its argument structure. In example (11), the agent BOY takes the subject function and presumably its position in the structure. Therefore, the ground cannot be the subject. At the same time, BOY as a subject marks the left periphery of the clause. Since the ground TABLE precedes the subject, it obviously occupies the left periphery of the clause.

(10)



\_\_\_\_\_ re

THREE BOOK

CL(C)<sub>A</sub>

APPLE

BE-LOCATED(5-B)<sub>A</sub>

‘There is an/the apple on three books.’

(SZJ; m79)

(11)



\_\_\_\_\_ re

TABLE<sub>B</sub>

BOY

<sub>A</sub>PUT(5-B)<sub>B</sub>

APPLE

‘A boy put an apple under the table.’

(SZJ; m77)

Next, I go on to determine the exact type of movement that the ground undergoes. I focus on non-manuals that mark the ground constituent in the clause-initial position. In order to understand the phenomenon, I resort to the cross-linguistic research on information structure in sign languages. Let us examine the minimal pair in (12-13). The signer first signed (13), but immediately corrected herself and suggested (12) as a better variant. She later judged (13) as degraded, but not entirely ungrammatical. The crucial information is marked non-manually. Example (12) represents canonical locative constructions, and in canonical locative constructions, the ground (TABLE) is normally accompanied by

raised eyebrows. Ground is manually marked by raised eyebrows, presumably because it occupies a non-argumental position within the left periphery. This is further confirmed by a sharp break in signing after the ground is produced. This break is marked by a pause in signing, a blink of the eyes, and the eyebrows suddenly and explicitly returning to the neutral position. Referring to cross-linguistic literature on sign language prosody (Nespor and Sandler 1999; Herrmann 2010; Pfau and Quer 2010; Sandler 2012; Kimmelman 2014; Kimmelman and Pfau 2016), I assume that these markings signal prosodic breaks – however, note that SZJ prosody has not been researched at all. I now attempt to determine what kind of movement triggered the displacement of the ground to the left periphery of the sentence. Compare the non-manual behaviour of the figure in examples (12) and (13). In (12), the figure is found in its canonical position with respect to the ground and predicate sign. It is non-manually unmarked. In (13), the figure is not found in its canonical position with respect to the ground and predicate sign: it precedes the ground and is non-manually marked with raised eyebrows. As a result, in (13), both the figure and the ground are marked with the same non-manuals. I assume that they underwent the same type of movement. Since we have already seen in (12) that the ground is fronted through non-argumental movement, I assume that the same holds true for the figure in example (13). However, are the constituents displaced in this movement topicalised or focalised (they are not wh-fronted since the force of the sentences is declarative)?

- (12)  $\overline{\text{IX}}_A \text{TABLE}_A \overline{\text{IX}}_A \text{CL}(\text{B}) + \text{BASKET}_B \text{BE-LOCATED}(\text{A})_A$  (SZJ; m32b)  
 ‘There is a basket on the table.’
- (13)  $\overline{\text{BASKET}}_A \overline{\text{CL}}(\text{B})_B \overline{\text{IX}}_B \overline{\text{TABLE}}_B \text{BE-LOCATED}(\text{A})_B$  (SZJ; m32a)  
 ‘There is a basket on the table.’

Rizzi (1997) and a significant body of literature covering various unrelated oral and sign languages have demonstrated that there can only be one focus interpreted in a sentence – contrary to topicalised constituents that have two positions reserved within the left periphery of the clause. Therefore, it seems reasonable to assume that, in (13), the ground and the figure raise each to the specifier position of two distinct topic projections. Furthermore, in both (12) and (13), the ground is endowed with all three of the characteristics that Kimmelman and Pfau (2015) identify as significant for topic constituents in sign languages: (i) as far as the word order is concerned, topics tend to be fronted, (ii) as far as syntactic marking is concerned, topics tend to be marked with raised eyebrows, and (iii) as far as prosody is concerned, topics tend to be followed by a prosodic break. These characteristics are detected in various unrelated sign languages, among others in American (Aarons 1994; Todd 2008), Finnish (Jantunen 2007), Hong Kong (Sze 2008, 2011), Israeli (Rosenstein 2001: ISL), Russian Sign Language (Kimmelman 2012), and in Sign Language of the Netherlands (Coerts 1992; Crasborn et al. 2009). The ground constituents in SZJ locative constructions also display these

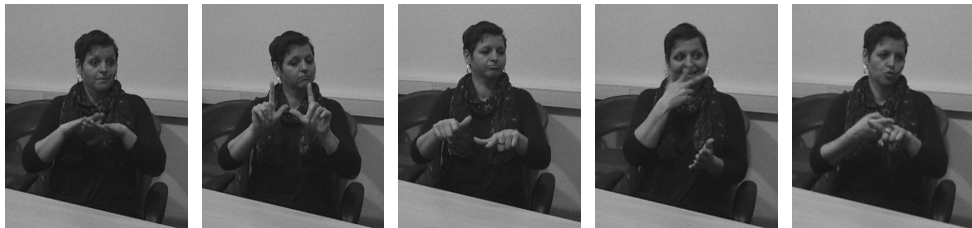
characteristics. Therefore, they are good candidates for topics. However, since topicalisation and the left periphery in general have not been researched in SZJ, the analysis proposed is still pending and its soundness awaits further evidence.

#### 4.2 Figure

In this subsection, I apply the distributivity morpheme test to the figure in order to confirm it as an internal argument of SZJ locative constructions. The distributivity morpheme repeats the verb, so that “each repetition of the verb has a start or an end point that is progressively further along the arc of the sweep” in order to “convey the information that the action was performed with respect to each member of the set of entities constituting the subject or object argument” (MacLaughlin et al. 2000: 85-86). According to the same authors, the verb is normally repeated three times, regardless of the number of elements in the set. According to Newman (2012), the distributivity morpheme quantifies over the subject of intransitives or over the object of transitives. This is roughly the same distribution that (i) Pavlič (2016) suggests for the distributive morpheme in SZJ and that (ii) Benedicto and Brentari (2004) suggest for the distributive morpheme in American Sign Language. Since in both languages distributive morphemes only attach to those verbs that license an internal argument, the distributive morpheme may be used as a test to verify the presence of an internal argument in these languages.

Let me return to the “vessel examples” above. In (4) and (5), a total of six repetitions of the verb are divided into two sets, so that the verb is repeated three times “progressively further along” the right bank and three times “progressively further along” the left bank. These six repetitions form two distributive morphemes. Since distributive morphemes attach exclusively to predicates that license an internal argument, I conclude that the figure functions as an internal argument in classifier-predicate examples (4), (5), and (6). A similar pattern may be observed in SZJ locative constructions with a lexical verb, such as LAY in (7). Again, according to the grammaticality judgments of my informants, the predicate may reduplicate (14). Thus, according to the distributivity-morpheme test, figures are base generated in the internal argument position.

(14)



————re

TABLE

LAY+DM

THREE

POTATOE

‘Three potatoes are lying on a/the table.’

(SZJ; loc-gjt4)

Finally, there is also the difference in word order between classifier and non-classifier predicates with respect to the figure. Is it the verb that moves in one case, but not in the other? Or is it actually the figure that moves from its base position with classifier predicates, but not with lexical predicates? Indeed, the effect of a classifier predicate on the word order in transitive and locative sentences is a well-known phenomenon in sign language linguistics, but in these two environments, the constituents themselves do not necessarily provide enough information about their relative position in the syntactic structure. Independent and at the same time indirect evidence derives from SZJ ditransitives (Pavlič 2016), in which the classifier predicate is found between the direct and the indirect object, yielding a non-basic  $SO_dVO_i$  word order (as compared to basic  $SVO_dO_i$ , displayed in SZJ ditransitives with a non-classifier predicate). This position seems to suggest that, in SZJ, classifier predicates do not move to higher sentential projections. In future research on SZJ, this tentative explanation must be confirmed by standard dominance tests, such as quantifier scope or anaphora binding.

## 5. CONCLUSION

In this research, I examined the morpho-syntactic properties of the predicates in locative constructions of Slovenian Sign Language (SZJ), as well as the syntactic functions of the figure and ground constituents. This enabled me to determine the surface word orders that are attested in the examples I provided. The ground appears as the first constituent in SZJ locative constructions, regardless of the type of predicate. The type of predicate, in turn, does influence the word order. I assume that the classifier predicate remains in situ, yielding a ground-figure-predicate (OSV) word order, while the lexical verb moves to higher sentential projections, yielding a ground-predicate-figure (OVS) word order. This answers research questions RQ1 – RQ3:

- A1 SZJ encodes locative information by modulating a spatially-agreeing predicate.
- A2 In SZJ locative constructions, a classifier predicate, a lexicalised classifier predicate, and a lexical verb may all encode a predicate.
- A3 The basic word order in SZJ locative constructions is OSV when a classifier predicate is used and is OVS when a lexicalised classifier predicate or a lexical verb is used.

## References

- AARONS, Debra (1994) *Aspects of the Syntax of American Sign Language*. Ph.D. Thesis. Boston University, Boston.
- BENEDICTO, Elena/Diane BRENTARI (2004) “Where Did All the Arguments Go? Argument-Changing Properties of Classifiers in ASL.” *Natural Language and Linguistic Theory* 22/4, 743–810.

- COERTS, Jane (1992) *Nonmanual Grammatical Markers: An Analysis of Interrogatives, Negations and Topicalisations in Sign Language of the Netherlands*. Ph.D. Thesis. University of Amsterdam, Amsterdam.
- COERTS, Jane (1994) "Constituent Order in Sign Language of the Netherlands and the Functions of Orientations." In: I. Ahlgren/B. Bergman/M. Brennan (eds), *Perspectives on Sign Language Structure*. Durham: ISLA, 69–88.
- COMRIE, Bernard (1989) *Language Universals and Linguistic Typology*. Oxford: Blackwell.
- CRASBORN, Onno/Els V. der KOOIJ/Johan ROS/Helen de HOOP (2009) "Topic Agreement in NGT (Sign Language of the Netherlands)." *The Linguistic Review* 26, 355–370.
- CULICOVER, Peter/Robert LEVINE (2001) "Stylistic Inversion in English: A Reconsideration." *Natural Language & Linguistic Theory* 19, 283–310.
- JANTUNEN, Tommi (2007) "On Topic in Finnish Sign Language." Unpublished Ms.
- JOHNSTON, Trevor/Myriam VERMEERBERGEN/Adam SCHEMBRI/Lorraine LEESON (2007) "Real Data are Messy: Considering Cross-Linguistic Analysis of Constituent Ordering in Auslan, VGT, and ISL." In: P. Perniss/R. Pfau/M. Steinbach (eds), *Visible Variation: Comparative Studies on Sign Language Structure*. Berlin: Mouton de Gruyter, 163–205.
- HERRMANN, Annika (2010) "The Interaction of Eye Blinks and Other Prosodic Cues in German Sign Language." *Sign Language and Linguistics* 13/1, 3–39.
- KIMMELMAN, Vadim (2012) "Word Order in Russian Sign Language." *Linguistics in Amsterdam*, 1–56.
- KIMMELMAN, Vadim (2014) *Information Structure in Russian Sign Language and Sign Language of the Netherlands*. Amsterdam: Amsterdam Center for Language and Communication (ACLC).
- KIMMELMAN, Vadim/Roland PFAU (2016) "Information Structure in Sign Languages." In: C. Fery/S. Ishihara (eds), *The Oxford Handbook on Information Structure*. Oxford: Oxford University Press, 814–835.
- LAUDANNA, Alessandro (1987) "Ordine dei segni nella frase." In: V. Volterra (ed.), *La Lingua Italiana dei Segni*. Bologna: Il Mulino, 211–230.
- LAUDANNA, Alessandro/Virginia VOLTERRA (1991) "Order of Words, Sign and Gestures: A First Comparison." *Applied Psycholinguistics* 12, 135–150.
- LIDDELL, Scott K. (1980) *American Sign Language Syntax*. Hague: Mouton Publishers.
- MACLAUGHLIN, Dawn/Carol NEIDLE/Benjamin BAHAN/Robert. G. LEE (2000) "Morphological Inflections and Syntactic Representations of Person and Number in ASL." *Recherches linguistiques de Vincennes* 29, 73–100.
- MATHUR, Gaurav/Christian RATHMANN (2012) "Verb Agreement." In: R. Pfau/M. Steinbach/B. Woll (eds), *Sign Languages*. Berlin/Boston: Mouton De Gruyter, 136–157.
- MILKOVIĆ, Marina/Sandra BRADARIĆ-JONČIĆ/Ronie B. WILBUR (2007) "Information Status and Word Order in Croatian Sign Language." *Clinical Linguistics and Phonetics* 21/11–12, 1007–1017.

- NESPOR Marina/Wendy SANDLER (1999) “Prosody in Israeli Sign Language.” *Language and Speech: Special Issue on Prosody in Spoken and Signed Languages* 42 (2/3), 143–176.
- NEWMAN, Paul (2012) “Pluractional Verbs: An Introduction.” In: P.C. Hofherr/B. Laca (eds), *Verbal Plurality and Distributivity*. De Gruyter, 185–211.
- PADDEN, Carol. (1983) *Interaction of Morphology and Syntax in American Sign Language*. Ph.D. Thesis, University of California, San Diego.
- PAVLIČ, Matic (2016) *The Word Order Parameter in Slovenian Sign Language: Transitive, Ditransitive, Classifier and Locative Constructions*. Ph.D. thesis. Ca’ Foscari, Venice.
- PERNISS, Pamela/Roland PFAU/Markus STEINBACH (2007) “Can’t You See the Difference? Sources of Variation in Sign Language Structure.” In: P. Perniss/R. Pfau/M. Steinbach (eds), *Visible Variation: Comparative Studies on Sign Language Structure*. Berlin: Mouton de Gruyter, 1–34.
- PFAU, Roland/Joseph QUER (2010) “Nonmanuals: Their Grammatical and Prosodic Roles.” In: D. Brentari (ed.), *Cambridge Language Surveys: Sign Languages*. Cambridge: Cambridge University Press, 381–402.
- DE QUADROS, Ronice Müller/Joseph QUER (2008) “Back to Back(wards) and Moving On: On Agreement, Auxiliaries and Verb Classes in Sign Languages.” In: R.M. De Quadros (ed.), *Sign Languages: Spinning and Unraveling the Past, Present and Future*. Petrópolis/RJ, Brazil: Editora Arara Azul, 530–551.
- RIZZI, Luigi (1997) “The Fine Structure of the Left Periphery.” In: L. Haegeman (ed.), *Elements of Grammar*. Dordrecht: Kluwer, 281–337.
- ROSENSTEIN, Ofra. (2001) *ISL as a Topic-Prominent Language*. Master’s Thesis. University of Haifa.
- SANDLER, Wendy (2012) “Visual Prosody.” In: R. Pfau/M. Steinbach/B. Woll (eds), *Sign Language*. Berlin/Boston: Mouton De Gruyter, 50–76.
- SZE, Felix. Y. B. (2008) *Topic Constructions in Hong Kong Sign Language*. Ph.D. Thesis. University of Bristol.
- SZE, Felix. Y. B. (2011) “Nonmanual Markings for Topic Constructions in Hong Kong Sign Language.” *Sign Language & Linguistics* 14/1, 115–147.
- TALMY, Leonard (2000) *Toward a Cognitive Semantics*. Cambridge, MA: MIT Press.
- TODD, Peyton (2008) “ASL Topics Revisited.” *Sign Language & Linguistics* 11/2, 184–239.
- VOLTERRA, Virginia/Alessandro LAUDANNA/Serena CORAZZA/Elena RADUTZKY/Francesco NATALE (1984) “Italian Sign Language: The Order of Elements in the Declarative Sentence.” In: F. Loncke/P. Boyes-Braem/Y. Lebrun (eds), *Recent Research on European Sign Languages*. Lisse: Swets and Zeitlinger, 19–48.



Summary  
SIGN ORDER IN SLOVENIAN SIGN LANGUAGE  
LOCATIVE CONSTRUCTIONS

In both sign and spoken languages, locative relations tend to be encoded within constructions that display the non-basic word/sign order. In addition, in such an environment, sign languages habitually use a distinct predicate type – a classifier predicate – which may independently affect the order of constituents in the sentence. In this paper, I present Slovenian Sign Language (SZJ) locative constructions, in which (i) the argument that enables spatial anchoring (“ground”) precedes both the argument that requires spatial anchoring (“figure”) and the predicate. At the same time, (ii) the relative order of the figure with respect to the predicate depends on the type of predicate employed: a non-classifier predicate precedes the figure, while a classifier predicate only comes after the figure.

**Keywords:** locative construction, locative adposition, figure and ground, classifier predicate, Slovenian Sign Language

Povzetek  
ZNAKOVNI RED V KRAJEVNIH IZRAZIH SLOVENSKEGA  
ZNAKOVNEGA JEZIKA

V krajevnihi stavkih sestavniki pogosto sledijo besednemu/znakovnemu redu, ki ni v skladu z osnovnim besednim/znakovnim redom danega govornega oziroma znakovnega jezika. Poleg tega je za znakovne jezike značilno, da se v tem skladenjskem okolju pojavlja poseben tip predikata (klasifikatorski predikat), ki že sam zase lahko vpliva na zaporedje sestavnikov v stavku. V članku predstavljam krajevne stavke slovenskega znakovnega jezika (SZJ). V njih je (i) argument, ki predstavlja referenčno točko (‘podlaga’), umeščen pred argument, ki potrebuje referenčno točko (‘lik’), in predikat; medtem ko je (ii) zaporedje lika in predikata odvisno od vrste predikata: ne-klasifikatorski predikat se umešča pred lik, klasifikatorski predikat pa za lik.

**Ključne besede:** krajevni stavki, krajevni predlogi, lik in podlaga, klasifikatorski predikati, slovenski znakovni jezik



## TYPES AND DEGREES OF VOWEL NEUTRALITY

### 1. INTRODUCTION

Traditionally, the distinction between neutral and harmonic behaviour in vowel harmony has been considered a categorical property of (sets of) vowels, which uniquely identifies each vowel of a harmony system as a member of one of two non-overlapping sets, i.e., the set of harmonic vowels and the set of neutral vowels<sup>1</sup> (e.g., van der Hulst and van de Weijer 1995). In this paper we focus on the ways in which neutrality is realised in suffixed forms and we argue that neutrality is not categorical. The graduality of neutrality has been discussed or suggested in the literature before (e.g., Anderson 1980) but has never been given a general explicit characterisation (as opposed to the degree of harmony, cf. Sanders and Harrison 2012; Alderete and Finley in press; also see section 3 for a discussion of the difference). We argue here that neutrality can manifest itself in different ways in different harmony systems since it derives from the ability of vowels to be involved in more than one pattern of neutrality, and a given pattern (i) may or may not be present in a system, (ii) may or may not combine with the other patterns, and (iii) may or may not be subject to variation – all of which influence the degree of neutrality. Based on this, we set up a scale of neutrality and offer a tentative quantification of the neutrality of harmony systems. We also show that in languages with more than one neutral vowel, the different neutral vowels may be of different types and degrees of neutrality. In the paper we only consider root/stem controlled front/back harmony systems with affix harmony where neutral vowels are phonetically front.<sup>2</sup> For simplicity's sake, we identify affixes with suffixes (although everything we say about harmony in suffixes would apply to prefixes as well).

---

\* rebrus@nytud.hu

\*\* tork@nytud.hu

1 In some languages, a condition is added which permits the neutrality/harmonicity of a given vowel to be specific to some prosodic-position or morphological domain (Kiparsky and Pajusalu 2003). Here we abstract away from these positional or domain-specific differences in the harmonicity/neutrality of one and the same vowel.

2 This is typically the case in front/back harmony.



## 2. PATTERNS OF NEUTRALITY

In a harmonic system, vowels are biased to co-occur in such a way that they agree in the value of a designated feature (the harmonic feature) within a morphologically and/or phonologically circumscribed domain. In a harmony system, vowel neutrality can be identified as the *lack of such bias*. This can manifest itself in three different ways: (a) a suffixal neutral vowel is unbiased as a ‘target’ in that it can systematically co-occur with vowels (potential triggers in the stem) of either harmonic value: this pattern of neutrality occurs when a stem is affixed with an *invariant suffix*. A neutral vowel in a root or a suffix is unbiased as a “trigger” in that (b) it does not modify the harmonic bias imposed on the target by other harmonic sources: this pattern of neutrality consists in *transparency* to the propagation of harmonic features; and (c) it does not impose its own harmony requirement on other vowels: this pattern of neutrality is *anti-harmony* requiring suffixation with suffix alternants that have a harmonic value opposite to that of the neutral vowel.<sup>3</sup> These possibilities<sup>4</sup> are summarised in Figure (1) below for front/back harmony, where neutral vowels are phonetically front (henceforward, we employ the following abbreviations and conventions: B/F/N: back/front/neutral vowels; [ ]: stem boundaries; consonants are left unindicated in formulaic descriptions, e.g., BN = C<sub>0</sub>BC<sub>0</sub>NC<sub>0</sub>).

### (1) Patterns of neutrality

N is unbiased as a:	name of the pattern	in formula
a. target	invariant suffixes (IV)	[[B]N] & [F]N]
b. trigger wrt. to another source	transparency (TP)	[[BN]B] & [[FN]F]
c. trigger as a source	anti-harmony (AH)	[[N]B]

The patterns (1a-b) above involve both harmonic values: (a) invariant suffixes necessarily occur after *both* front and back stems, and (b) transparency typically consists in the neutral vowel’s lack of bias while passing on *both* the front and the back bias of the preceding harmonic vowel.<sup>5</sup> Anti-harmony (1c), however, *only* involves a harmonic

3 A reviewer points out that lack of harmony (presumably in the form of hesitation) could in principle be the realization of a neutral vowel’s lack of bias as a trigger. This is a valid point, but, interestingly, this state of affairs seems to be unattested, cf. Aoki (1968). Note that our use of the term *antiharmony* is different from Sanders and Harrison (2012), who use it to refer to preference for disharmony globally in a system.

4 Another possible (and attested) pattern of neutrality is one that is completely contained within the monomorphemic form. This is the “mixed” root pattern where the neutral/unbiased character of a vowel manifests itself in its ability to freely co-occur with any other vowel of the system within a root. We disregard this pattern in this paper since we are focusing on neutrality in affix harmony. For a discussion of mixed stems in Hungarian and their relevance to neutrality, cf. Törkenczy (2011) and Törkenczy et al. (2013).

5 It is a definitive property of the IV neutrality pattern (1a) that an allomorph with the *same* vowel occurs after both front harmonic and back harmonic stems (this is indicated by the symbol “&” in

value in the suffix which is the opposite of the phonetic value of the (neutral vowel-) stem, i.e., [[N]B], but not necessarily [[N]F].<sup>6</sup>

### 3. TYPOLOGY

Each of the patterns in (1) may be present in or absent from a given harmony system (cf. Kiparsky and Pajusalu 2003).<sup>7</sup> Thus, neutrality can manifest itself in more than one way and the different patterns may or may not co-occur in a given system. Therefore, it is possible to define *degrees of neutrality* and set up a *classification of systems*<sup>8</sup> according to the types of neutrality they permit. Naturally, other classifications are also possible, e.g., ones that are based on a measure of harmony assessed “globally” in a system. Some recent proposals (Sanders and Harrison 2012; Alderete and Finley in press) measure the degree of global (dis)harmony of a system independently of any specific neutrality pattern (such as transparency/opacity, antiharmony etc.). In our approach the degree of neutrality of a (sub)system depends on the “strength” of the neutrality patterns (IV, TP and AH) measured in relative word-type frequency (the type frequency of word-forms (not lemmas or tokens)). This is different from these global approaches, which characterise systems by measuring the degree of global (dis)harmony in word-types (Sanders and Harrison 2012) or stem-types (Alderete and Finley in press) within a system. A crucial difference is that disharmony (co-occurring vowels that disagree in their specifications for the harmonic feature) is not the same as the neutrality defined in this paper (involvement of a vowel in (some of) the neutrality patterns IV, TP, AH) because (a) a neutrality pattern does not necessarily result in disharmony (antiharmony [[N]B] is always disharmonic but the other neutrality patterns are not necessarily so, e.g., in TP, transparency to backness is disharmonic, [[BN]B], but transparency to frontness is not [[FN]F]), and (b) disharmony is not necessarily associated with the neutrality patterns we examine here (e.g., root-internal [NB] is disharmonic, but does not realise IV, TP or AH).

---

the relevant formula in the rightmost column). However, in the TP neutrality pattern (1b), transparency to the *frontness* of a trigger vowel and transparency to the *backness* of a trigger vowel are logically independent and may be assessed independently. Indeed, according to Kiparsky and Pajusalu (2003), there is a rare type of harmony system in which transparency to backness occurs but transparency to frontness does not ([[BN]B] & [[FN]B]), while “anti-transparency” ([[BN]F] & [[FN]B]) is unattested in harmony systems. For the sake of simplicity, we do not distinguish these two kinds of transparency in this paper, and pattern (1b) covers both: [[BN]B] & [[FN]F].

6 This also means that in a system with *total* anti-harmony, all stems with a neutral vowel “trigger” behave like [[N]B] and no stems behave like [[N]F]. Practically, most anti-harmonic systems show both; see section 4 for variation in AH.

7 Although they cannot combine arbitrarily, see Kiparsky and Pajusalu (2003). AH entails TP for backness. There is no similar constraint on the occurrence of invariant suffixes, so theoretically IV can combine freely with the two other patterns of neutrality. In this section we examine only those systems where neutrality involves invariant suffixes (i.e., TP entails IV), but in section 4 the Hungarian examples involve variation for all the three patterns, including IV.

8 Or subsystems; see section 5.

In table (2) below we show four language types (2o-iii), which exhibit different types of neutrality of Ns occurring in a harmony system. Since we have defined neutrality as the *lack* of bias, the different properties of the patterns (1a-b) vs. (1c) discussed above affect the degree of neutrality of a system differently. The presence of IV and/or TP truly means a lack of bias since a neutral vowel can occur after a vowel of either harmonic value (IV) and a vowel of either harmonic value can occur after a neutral vowel (TP). By contrast, total AH in a system *is* a kind of bias because only B can occur after N in a suffix. Therefore, none of the language types (2o-iii) in (2) have maximal neutrality, because neither the ones without AH (2o-ii), nor the ones with AH (2iii) have ‘ideal’ neutral properties, i.e., a complete lack of bias in this respect.

(2) Some language types based on neutrality patterns (1a-c)

language types – examples	a. IV	b. TP	c. AH	neutrality (lack of bias)
o. no neutral vowel – Turkish	–	–	–	no
i. only IV, no TP/AH – E Khanty /i/	+	–	–	low
ii. IV & TP, no AH – Finnish /i/, /e/	+	+	–	high
iii. IV & TP & AH – Uyghur /i/, /e/	+	+	+	high

Note that (2) is a descriptive kind of typology rather than a “factorial” one. It shows a sample of the possible types of harmony systems representing various degrees of neutrality rather than limiting what is a possible system with (or without) neutrality; e.g., a (front/back) system with AH but without TP does not occur to the best of our knowledge and is arguably impossible (cf. Kiparsky and Pajusalu 2003; Rebrus and Törkenczy 2015).<sup>9</sup> Nevertheless, it could be added to Table 2 where it would represent a system of intermediate neutrality just like (2ii) or (2iii), which *are* attested.

#### 4. VARIATION

Patterns (1a-c) may be subject to variation:<sup>10</sup> (a) a given vowel may occur in both invariant and harmonically alternating suffixes in a system, (b) transparency–opacity may be variable [[BN]B/F], and/or (c) anti-harmony may be variable [[N]B/F]. This yields further language types, possibly of different degrees of neutrality. As an example, consider a harmony system with invariant neutral vowel suffixes (both [[B]N] and [[F]N] occur), variable transparency (both [[BN]B] and [[BN]F] (and also [[FN]

9 See Rebrus and Törkenczy (2015ab) for a discussion and a principled explanation that attributes the restrictions on patterns to monotonicity.

10 Here we do not distinguish between lexical variation and vacillation and use the symbol “/” to denote variation (of either kind), see Rebrus and Törkenczy 2015b.

F]) occur), and without antiharmony ([N]B does not occur). Such a system (see 3ii' below) would be somewhere between (3i=2i) and (3ii=2ii) in the tentative scale in (2) because here N does have some bias in transparency since [[BN]F] also occurs with some probability (indicated by “+/-” in (3)). A system with variable IV (in which there is lexical variation such that some N-vowel suffixes are invariable while others alternate harmonically) but without TP or AH is even less neutral than the invariable system in (3i=2i) because in such a system, some suffixes with N are targets of harmony. A system of maximal neutrality (3iii' below) has (in addition to invariant suffixes and invariable transparency) *variable* antiharmony since in this case the neutral vowel of an all-neutral root imposes no bias at all on the suffix (i.e., both B and F can occur in the suffix<sup>11</sup>). An approximate neutrality scale of harmony types that includes these variable systems as well is shown in (3).

(3) Degrees of neutrality in types of harmony systems with and without variation

language types	a. IV	b. TP	c. AH	neutrality
o. no neutral vowel	–	–	–	no
i'. <i>variable</i> IV, no TP/AH	+/-	–	–	very low
i. only IV, no TP/AH	+	–	–	low
ii'. IV, <i>variable</i> TP, no AH	+	+/-	–	intermediate
ii. IV & TP, no AH	+	+	–	high
iii. IV & TP & AH	+	+	+	high
iii'. IV & TP & <i>variable</i> AH	+	+	+/-	<b>maximal</b>

## 5. GRADUALITY

More than one neutral vowel occurs in some harmony systems. In such a system neutrality may be *homogeneous*, i.e., all neutral vowels behave the same way: this is the case of /i/ and /e/ in Finnish and Uyghur (cf. Anderson 1980; Vaux 2000) – see (2ii,iii); or it can be *non-homogeneous* when the different neutral vowels show different degrees of neutrality because they are involved in the patterns (3i'-iii') differently. We consider languages of this kind *polysystemic*, i.e., they contain vowel harmony subsystems of more than one type such that the different types are specific to different (groups of) neutral vowels. Front/back harmony in Hungarian is an example where this is referred to as the *height effect* (e.g., Hayes and Cziráky Londe 2006), which means that the higher a front unrounded vowel is, the more neutrally it behaves. This is illustrated in (4) below, where rows contain the different Hungarian front unrounded vowels /i i: e: ε/ and the columns show the patterns of neutrality. With the exception of /i:/ all the neutral vowels in Hungarian show variation in IV in addition to variation in TP and/or AH (we discuss the Hungarian data in more detail below).

11 Assuming that they both occur with the same probability.

(4) Hungarian: differences in the neutrality of neutral vowels

language types	a. IV	b. TP	c. AH	neutrality
ii''. <i>variable</i> IV/TP, no AH – Hungarian / $\epsilon$ /	+/-	+/-	-	low
iii''. <i>variable</i> IV/TP/AH – Hungarian / $e$ /	+/-	+/-	+/-	intermediate
iii'. TP & <i>variable</i> IV/AH – Hungarian / $i$ /	+/-	+	+/-	high
iii'. IV & TP, <i>variable</i> AH – Hungarian / $i$ /	+	+	+/-	maximal

Note that – while the systems in the sample of possible types shown in (3i'-iii') can be naturally and uniquely arranged along a scale of neutrality ranging from minimally neutral to maximally neutral – it is not possible to arrange all the *theoretically possible* types in a similar scale. For instance, it is not possible to determine if a system with variable IV and invariable TP (<+/-, +, ->) is more or less neutral than its “mirror image”, i.e., a system with invariable IV and variable TP (<+, +/-, ->). Similarly, it is not possible to determine the difference in neutrality between (3i) <+, -, -> and (4ii'') <+/-, +/-, ->. <sup>12</sup> This in itself is a motivation for characterizing or calculating the neutrality of a vowel with a value whose assignment is consistent across all vowels.

We present here a tentative quantification of the neutrality of Hungarian vowels which fall into several types based on their behaviour in the Hungarian front/back harmony system (summarised in (5)). The numbers in (5) below and in the text indicate the approximate degree of neutrality on a 5-point scale from [0] to [4] ([0]=non-neutral, [4]=totally neutral, [1,2,3]=variably neutral) depending on their involvement in the three patterns of neutrality (IV, TP and AH) that occur in the Hungarian system. It must be pointed out here that there is no *theoretical* reason why neutrality should be assessed on a 5-point scale rather than a scale of fewer or more than five degrees of neutrality; the motivation is practical: qualitatively, 5 degrees of neutrality behaviour characterise the Hungarian system. The vowels traditionally considered as neutral /i, i:, e:,  $\epsilon$ /<sup>13</sup> are involved in these neutrality patterns in the following way.

(a) Occurrence in invariant suffixes: long /i:/ only occurs in invariant suffixes (e.g., verb-forming **-it**, **ton-it** ‘teach’) and is therefore completely neutral in this respect [4]; short /i/ mostly occurs in invariant suffixes (e.g., Terminative **-ig**, **haz-ig** ‘up to the house’) but it does occur in an alternating one (Pres.Sg3.Def **-i** ~ **jo**, **dob-jo** ‘throw’, **yt-i** ‘hit’); invariant : alternating ratio=11:1, so it is less neutral [3]. / $\epsilon$ / is frequent in

12 This is not true of the Hungarian data shown in (4), which, again, represent a subset of the possible types and happen to be such that they can be arranged in a scale of neutrality in a natural and unique way, as shown in the last column of (4). The reason is that in this sample it is never the case that a given vowel  $V_a$  is more neutral than another  $V_b$  with respect to one pattern of neutrality while  $V_b$  is more neutral than  $V_a$  with respect to another pattern.

13 The neutrality of / $\epsilon$ / has been debated in the literature of Hungarian vowel harmony (see Törkenczy 2011 for an overview) but / $\epsilon$ / can be given comfortable place as soon as we abandon the view that the distinction between neutral and non-neutral is categorical (which is our view here and see also Hayes and Cziráky Londe 2006).

both invariant suffixes (e.g., Causal **-ert**, **ha:z-ert** ‘for the house’) and alternating suffixes (e.g., Adessive **-ne:l** ~ **na:l**, **kert-ne:l** ‘at the garden’, **ha:z-na:l** ‘at the house’); invariant : alternating=7:10<sup>14</sup>, and thus it is less neutral than /i/ [2]. /ε/ is frequent in alternating suffixes (e.g., Inessive **-ben** ~ **bön**, **kert-ben** ‘in the garden’, **ha:z-bön** ‘in the house’), but it occurs in invariant ones only in a handful of special diminutive constructions (i.e., Diminutive **-es**, **kor-es** ‘Charley’); invariant : alternating=2:50, so it is even less neutral [1].

(b) Variability in transparency: /i, i:/ are always transparent in all roots (e.g., **ko:fi-nök**/\***ko:fi-nek** ‘car-DAT’, **po:pi:r-nök**/\***po:pi:r-nek** ‘paper-DAT’); thus, they are completely neutral [4] in this respect. /e:/ is usually transparent with some root-specific lexical variation (invariable back-suffixed [Be:] stems e.g., **ka:ve:r-nök**/\***ka:ve:r-nek** ‘coffee-DAT’ are much more frequent than variably suffixed [Be:] stems e.g., **örze:n-nök**/**örze:n-nek** ‘arsenic-DAT’): thus, it is less neutral than /i, i:/ [3]. /ε/ shows a high degree of lexical variation and vacillation (variable suffixation, e.g., **ho:tel-nök**/**ho:tel-nek** ‘hotel-DAT’, is more frequent than invariable front suffixation, e.g., **ko:ntser:ne:k**/\***ko:ntser:ne-nök** ‘concert-DAT’ or mostly invariable back suffixation, e.g., **mö:sek-nök**/\***mö:sek-nek** ‘self-employed-DAT’): thus, it is weakly neutral in this respect: [2].

(c) Antiharmony: /i, i:/ frequently occur in anti-harmonic stems and almost all anti-harmonic stems have these vowels; therefore, they are highly neutral [4] in this respect. The vowel /e:/ very rarely occurs in antiharmonic stems (there are only two antiharmonic free stems and a few bound stems with /e:/, thus: [2]), and /ε/ practically does not, hence [0]. Table (5) shows these values for all the front vowels in Hungarian. We have also included front rounded vowels, which are non-neutral (i.e., they only occur in harmonically alternating suffixes, they are invariably opaque and do not occur in antiharmonic roots), to facilitate comparison with the neutral ones.

In the last column, we have given the average scores of the vowels in points and converted to percentages where 100% means complete neutrality and 0% means the total lack of neutrality (fully harmonic behaviour). These values represent the neutrality scores of the vowels. It can be seen in (5) that the neutrality scores of the vowels that are traditionally considered neutral fall within a rather wide range: from a 100% neutral /i:/ through “half-neutral” /e:/ (58%) down to the least neutral /ε/ (25%).

---

14 Stem final *a* and *e* lengthen to *a:* and *e:*, respectively, before suffixes by the productive process of Low Vowel Lengthening (see e.g., Siptár and Törkenczy 2000). This also affects suffixes that are *ε*-final when word-final, e.g., the possessive suffix **-(j)a/ε** – compare **fy:l-ε** ‘ear-POSS.3SG’ and **fy:l-ε-t** ‘ear-POSS.3SG-ACC’. If we also include these suffixes, then the invariant–alternating ratio changes to 7:13.

(5) Tentative estimation of the degree of neutrality for front vowels in H. (5-point scale: 0,...,4)

front vowels	a. IV	b. TP	c. AH	average (neutrality score)
high i: (= totally neutral)	4	4	4	4.0 (100%)
high i	3	4	4	3.7 (92 %)
high-mid e:	2	3	2	2.3 (58 %)
low-mid e	1	2	0	1.0 (25 %)
round y y: ø ø: (= non-neutral)	0	0	0	0.0 (0 %)

Note that the numbers in each of the columns (5abc) decrease<sup>15</sup> from top to bottom; thus, (5a), (5b) and (5c) each yield the same scale independently as the scale in the last column, which means that the height effect manifests itself in all the three patterns of neutrality. This is not a theoretical necessity, but a property of the Hungarian system, which is a sample of the theoretically possible types (it is possible to have a decreasing average score of neutrality while some of the patterns of neutrality do not show a decreasing scale).

The main problem with the quantification is that the scores for each of the neutrality patterns are given “impressionistically” and therefore a five-point *equidistant* scale [0,1,2,3,4] is not entirely justified. For instance, there is no reason why the difference between the TP of /i, i:/ [4] and the TP of /e:/ [3] should be identical with the difference between the TP of /e:/ [3] and the TP of /ε/ [2]. Also nothing guarantees that the “intermediate” degrees truly mean the same across patterns, i.e., for example /e:/ is truly equally neutral with respect to IV and AH (both of its scores are [2]).

A more realistic way of quantifying vowel neutrality is based on the *frequency ratios* of items realising these patterns. In order to develop such a quantification we have carried out a corpus study in Hungarian where we have made the following measurements and calculations for a frequency-based quantification of neutrality. We have used the *Szószablya* web corpus of the Hungarian language (Halácsy et al. 2004), which contains 541 million word tokens and 2.32 million word types.

We have calculated the TP neutrality ratio of the 4 neutral vowels in the following way: in each of the four classes of disyllabic [BN]-type stems, which differ in the neutral vowel ([Bi:], [Bi], [Be:], [Bε]) we have counted how many of the harmonically suffixed word-*types* contain a front suffix alternant ([[BN]F]) and how many of them contain a back suffix alternant ([[BN]B]).<sup>16</sup> Then, we have calculated the percentage of the back-suffixed word-types compared to the number of all the word types (F-suffixed and B-suffixed) in the given class, i.e., the following “backness” ratio:

15 In a weak sense that permits identical adjacent values.

16 We have counted types and not tokens – we do not consider how many tokens of a word type are found in the corpus, e.g., the types *fotelbõn* and *fotelben* ‘armchair-in.’ are both found in the corpus, but in our calculations it does not matter that the former is found 374 times and the latter 376 times since each type counts once.



$$(6) \quad p_{\text{TP}} = \frac{\text{freq}([\text{BN}]\text{B})}{\text{freq}([\text{BN}]\text{B}) + \text{freq}([\text{BN}]\text{F})}$$

In the case transparency,  $p$  defined above is used for the degree of neutrality. If it is 100%, then only B-suffixed word-types occur, showing that transparency (TP) is maximal, which means that the neutral vowel examined is completely neutral with respect to TP. It can be seen in column (b) in (8) below that /i:, i/ are maximally neutral ( $p=100\%$ ) and /e:/ is near-maximally neutral ( $p=96.8\%$ ) in this respect. Although opacity is more frequent than transparency in the case of /ε/, /ε/ behaves transparently in a quarter of the word-types ( $p=25.3\%$ ).

In order to quantify neutrality in antiharmony, we have examined the same relative frequencies of word-types, but this time those that containing monosyllabic [N] stems with the four neutral vowels [i:], [i], [e:], [ε]. Then we have converted the resulting backness ratios (a number between 0% and 100%), so that a backness ratio of exactly 50% (a hypothetical type of variation in which front and back suffixed word types occur with equal probability) corresponds to 100% neutrality and a backness ratio of 100% or 0% (no variation: complete antiharmony and complete lack of antiharmony, respectively) corresponds to 0% neutrality. We get the neutrality degree  $p$  in AH by taking the distance of a relative word type frequency (backness ratio) from 50% (the theoretical maximum of neutrality in AH), multiply it by two and calculate the distance of this number from 100%, as shown in (7), where  $r_{\text{AH}}$  is the backness ratio and  $p_{\text{AH}}$  is the degree of neutrality:

$$(7) \quad r_{\text{AH}} = \frac{\text{freq}([\text{N}]\text{B})}{\text{freq}([\text{N}]\text{B}) + \text{freq}([\text{N}]\text{F})} \quad p_{\text{AH}} = 1 - 2 \cdot |0.5 - r_{\text{AH}}|$$

The relative word-type frequencies (backness ratios  $r_{\text{AH}}$ ) are the following: /i:/: 51.8%, /i/: 10.4%, /e:/: 2.1%, and /ε/: 0.0%. The figures show that the number of harmonic and antiharmonic word-types with a long /i:/ in the monosyllabic root is approximately the same (in fact there are slightly more antiharmonic types than harmonic ones), and thus the neutrality of /i:/ with respect to antiharmony is close to the theoretical maximum ( $p=96.4\%$ ). The relative word-type frequency is much lower for /i/ (10.4%), i.e., only one word type with a root internal short /i/ out of ten is antiharmonic; therefore, the neutrality of /i/ is low ( $p=20.8\%$ ). The antiharmonicity values (expressed in backness ratios) of /e:/ and /ε/ are even lower, 2.1% and 0%, respectively, which corresponds to very low neutrality ratings with respect to antiharmony:  $p=4.2\%$  and  $p=0.0\%$ , respectively. This is shown in (8c).

Finally, consider the neutrality ratings with respect to invariance in suffixes. Here we have calculated the neutrality values using the statistics about the distribution of vowels in alternating and invariant suffixes given in the discussion of table (5) above. The neutrality rating of a neutral vowel with respect to invariance is the ratio of invariant (i.e., harmonically non-alternating) suffixes containing the given neutral vowel to all the suffixes (invariant or harmonically alternating) that contain it. This is shown



in (8a). All the suffixes that have long /i:/ are invariant, so it is maximally neutral with respect to invariance ( $p=100\%$ ). There is only one alternating suffix with short /i/ while all the others are invariant ( $p=91.7\%$ ). The other neutral vowels /e:/ and /ɛ/ are considerably less neutral with respect to invariance: only about a third of the suffixes with /e:/ are invariant ( $p=35.0\%$ ) and only two suffixes with /ɛ/ are invariant while a great majority of them alternate ( $p=3.8\%$ ).

(8) Neutrality scores (in percentage) calculated from type frequency ratios

front unrounded vowels	a. IV	b. TP	c. AH	average (neutrality score)	
				by frequency	in (5)
high i:	100	100	96.4	<b>98.8</b>	100
high i	91.7	100	20.8	<b>70.8</b>	92
high-mid e:	35.0	96.8	4.2	<b>45.3</b>	58
low-mid ɛ	3.8	25.3	0.0	<b>9.7</b>	25

## 6. CONCLUSION

As a conclusion let us examine what our study shows about (i) the relationship between the neutrality scores of the different vowels in a given pattern and on the average, and (ii) the relationship between the neutrality scores of one and the same vowel in the different neutrality patterns (a, b, c).

(i) The average neutrality scores of the neutral vowels are higher in (5) than in (8). This is due to the fact that the impressionistic five-point scale (5) is based on is not fine-grained enough. Nevertheless, the tendency of the scores in (5) and (8) is the same, i.e., they arrange the four neutral vowels in the same hierarchy of neutrality. The same relationship holds true of the average frequency-based neutrality scores of (8) (and (5)), and also of the scores of the vowels on the individual neutrality patterns, i.e., the neutrality scores decrease in each of the columns (8abc). This confirms the general view about the neutrality differences between the Hungarian neutral vowels (the height effect), which is typically either simply stipulated in studies (Siptár and Törkenczy 2000) or based on a single neutrality pattern, TP (Ringen and Kontra 1989; Hayes and Cziráky Londe 2006). It has been claimed in the literature that the height effect, i.e., the correlation between the increase of openness of a phonetically front neutral vowel and the decrease of neutrality is natural cross-linguistically (e.g., Anderson 1980; Kiparsky and Pajusalu 2003) and even phonetically motivated (in TP e.g., Beňuš 2005; Beňuš and Gafos 2007).

(ii) It can be seen in (8) that relationship between the neutrality scores for the three neutrality patterns is such that for every neutral vowel the neutrality score in transparency is always higher than the score in antiharmony and the score in invariance is always between transparency and antiharmony:

$$(9) \quad p_{TP} \geq p_{IV} \geq p_{AH}$$

This is graphically represented in Figure 1 below:

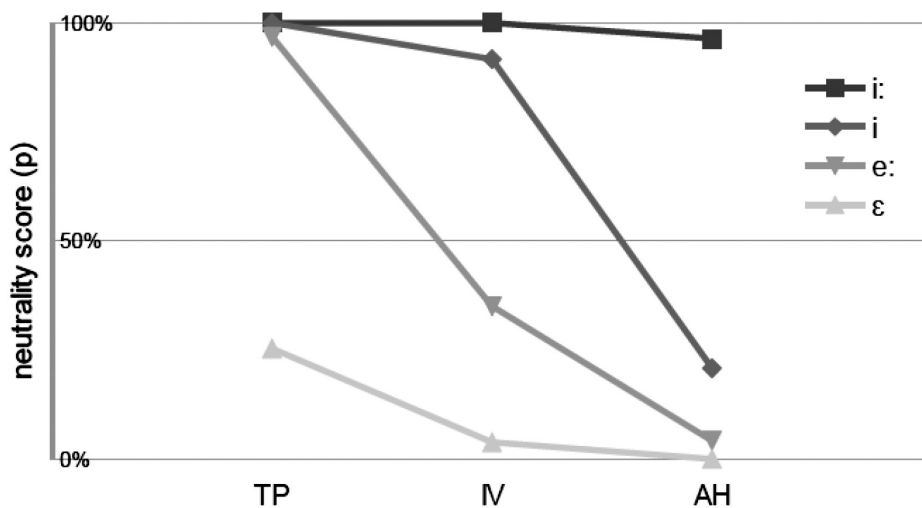


Figure 1. Neutrality score of different vowels by different patterns

It is reasonable to ask if this relationship in (9) (a) is just a parochial fact about the patterns Hungarian neutral vowels are involved in or one that holds universally/cross-linguistically, or (b) whether it has some explanation/motivation. Unfortunately, at the present state of our knowledge neither question can be given a clear answer. First, we know from studies of front/back harmony systems (cf. Kiparsky and Pajusalu 2003; Rebrus and Törkenczy 2015a) that the presence of AH in a system implies the presence of TP and that this can be generalised to variable TP and variable AH: the backness ratio of TP must be higher than the backness ratio of AH. This follows from the general (universal) principle of monotonicity that constrains (front/back) harmony systems (cf. Rebrus and Törkenczy 2015ab). However, the fact that  $r_{TP} \geq r_{AH}$  holds generally for backness ratios does not necessarily mean that  $p_{TP} \geq p_{AH}$  holds for neutrality scores as well, since the interpretation of neutrality (the relationship between a backness ratio and the corresponding neutrality score) is different for TP and AH (see sections 3 and 5 above). The placement of IV in between TP and AH universally is even more uncertain since we know less about it cross-linguistically: the available studies typically assume that it exists in the systems in which they examine TP (and possibly AH).<sup>17</sup> As

17 We know that there are systems in which harmony is limited to the root (e.g., Marash dialect of Armenian, cf. Vaux 1998). Trivially, in such a system, all affixes are invariant, thus there are systems with IV, but without TP and AH. However, these systems are irrelevant for us here since we are only interested in systems in which TP and AH, both of which manifest themselves in affix alternations, could in principle occur.

for question (b), i.e., the explanation/motivation of (9), we can look at the problem in two ways. On the one hand one could look at the trigger for potential disharmony and argue that in front/back harmony the occurrence of a harmonic value B is motivated in the environment of another B even if it is not strictly adjacent: BNB (as in TP), but it is not otherwise: NB (as in AH) or BN (as in IV). Therefore, it is to be expected that TP should be more frequent than AH or IV. This view groups TP (motivated) vs. AH, IV (unmotivated). On the other hand, it is also possible to look at potential disharmony as a function of the target. In the case of IV disharmony is motivated since by definition the suffix is invariably N, thus it cannot change since it does not alternate. By contrast, in TP and AH, disharmony is not motivated in this sense since the suffix is an alternating one and it could harmonise. Thus, one would expect IV to be more frequent than TP or AH. This view groups IV (motivated) vs. TP, AH, (unmotivated). A combination of the two views establishes AH as the least frequent, but does not determine a relationship between TP and IV. It must be noted that, again, this argument is about relative frequencies (backness ratios) rather than neutrality scores and thus, for the same reason as discussed above, does not really explain the relationship in (9).

This work has been supported by National Scientific Grant OTKA-104897 ‘Variation in Phonology’. We would like to thank the anonymous reviewer for drawing our attention to recent relevant literature.

## References

- ALDERETE, John/Sarah FINLEY (in press) “Gradient vowel harmony in Oceanic.” To appear in *Language and Linguistics* 17.
- AOKI, Haruo (1968) “Toward a Typology of Vowel Harmony.” *International Journal of American Linguistics* 34, 142–145.
- ANDERSON, L. (1980) “Using asymmetrical and gradient data in the study of vowel harmony.” In: R. Vago (ed.), *Issues in Vowel Harmony*. Amsterdam: John Benjamins, 271–340.
- BEŇUŠ, Stefan (2005) *Dynamics and transparency in vowel harmony*. PhD dissertation. New York University.
- BEŇUŠ, Stefan/Adamantios I. GAFOS (2007) “Articulatory characteristics of Hungarian ‘transparent’ vowels.” *Journal of Phonetics* 35, 271–300
- HALÁCSY, Péter/András KORNAI/László NÉMETH/András RUNG/István SZAKADÁT/Viktor TRÓN (2004) “Creating open language resources for Hungarian.” In: *Proceedings of Language Resources and Evaluation Conference (LREC04)*. LREC, 203–210.
- HAYES, Bruce/Zsuzsa CZIRÁKY LONDE (2006) “Stochastic phonological knowledge: The case of Hungarian vowel harmony.” *Phonology* 23, 59–104.
- HULST, Harry van der/Jeroen van de WEIJER (1995) “Vowel harmony.” In: John A. Goldsmith (ed.) 1995. *The handbook of phonological theory*. Cambridge, MA & Oxford: Blackwell, 495–534.

- KIPARSKY, Paul/K. PAJUSALU (2003) "Towards a typology of disharmony." *The Linguistic Review* 20, 217–241.
- REBRUS, Péter/Miklós TÖRKENCZY (2015a) "Monotonicity and the typology of front/back harmony." *Theoretical Linguistics* 41/1–2, 1–61.
- REBRUS, Péter/Miklós TÖRKENCZY (2015b) "The monotonic behaviour of language patterns." *Theoretical Linguistics* 41/3–4, 241–268.
- RINGEN, Catherine/Miklós KONTRA (1989) "Hungarian neutral vowels." *Lingua* 78, 181–191.
- SANDERS, Nathan/K. David HARRISON (2012) *Discovering new vowel harmony patterns using a pairwise statistical model*. Poster presented at the Twentieth Manchester Phonology Meeting, Manchester, UK, 24–26 May 2012. <http://www.swarthmore.edu/SocSci/dharris2/pdfs/sanders-harrison-mfm20-handout.pdf>
- SIPTÁR, Péter/Miklós TÖRKENCZY (2000) *The phonology of Hungarian*. Oxford: Oxford University Press.
- TÖRKENCZY, Miklós (2011) "Hungarian vowel harmony." In: M. van Oostendorp/C. J. Ewen/E. Hume/K. Rice (eds), *The Blackwell Companion to Phonology*. Malden, MA & Oxford: Wiley-Blackwell, 2963–2990.
- TÖRKENCZY, Miklós/Péter REBRUS/Péter SZIGETVÁRI (2013) "Harmony that cannot be represented." In: J. Brandtler/V. Molnár/C. Platzack (eds), *Approaches to Hungarian. Papers from the 2011 Lund Conference Volume 13*. Amsterdam/Philadelphia: John Benjamins, 229–252.
- VAUX, Bert (1998) *The Phonology of Armenian*. Oxford: Clarendon Press.
- VAUX, Bert (2000) "Hungarian vowel harmony. Disharmony and derived transparency in Uyghur vowel harmony." In: M. Hirotani/A. Coetzee/N. Hall/J.-Y. Kim (eds) *Proceedings of NELS* 30, 671–698.

## Summary

### TYPES AND DEGREES OF VOWEL NEUTRALITY

This paper argues that neutrality in a harmony system is a gradient property since it is due to a vowel's participation in different patterns that are considered to be indicators of neutral behaviour in harmony. We examine three of these patterns of neutrality (transparency, affixal invariance and antiharmony) and show that a scale of neutrality can be defined on the basis of these patterns (their occurrence and variability) and the neutrality of harmony systems can be characterized with reference to this scale. We describe a tentative quantification of neutrality and then develop an explicit measure of neutrality based on relative word type frequency. This explicit measure is applied to the behaviour of neutral vowels in Hungarian front/back harmony where the different neutral vowels represent different degrees of neutrality in all three neutrality patterns.

**Keywords:** vowel harmony, neutral vowels, gradience, Hungarian front/back harmony, variation

Povzetek  
VRSTE IN STOPNJE SAMOSTALNIŠKE NEVTRALNOSTI

Članek obravnava nevtralnost v sistemu harmonije kot stopnjevano lastnost, ki nastane zaradi udeležnosti samoglasnika v različnih vzorcih, ki so kazalci nevtralnega vedenja v harmoniji. V prispevku predstavimo tri takšne vzorce nevtralnosti (transparentnost, nespremenljivost pon in antiharmonijo) in pokažemo, da lahko na podlagi teh vzorcev (prek njihove pojavnosti in sprejemljivosti) določimo lestvico nevtralnosti, s katero lahko opišemo nevtralnost sistemov harmonije. Nevtralnost poskušamo tudi kvantificirati in nato razviti eksplicitno merjenje nevtralnosti, ki temelji na relativni pogostnosti tipov besed. Meritev uporabimo za nevtralne samoglasnike v harmoniji madžarskih sprednjih/zadnjih samoglasnikov, pri kateri različni nevtralni samoglasniki predstavljajo različne stopnje nevtralnosti v vseh treh vzorcih nevtralnosti.

**Ključne besede:** samoglasniška harmonija, nevtralni samoglasniki, stopnjevanost, harmonija sprednjih/zadnjih samoglasnikov v madžarščini, variacija



## ASYMMETRIES IN SUB-EXTRACTION OUT OF NP IN SLOVENIAN: A MAGNITUDE ESTIMATION STUDY\*\*\*\*

### 1. THE SUBJECT ISLAND IN SLOVENIAN

Golden (1995, 1996, 1997a: ch.8, 1997b) reports sentences in which a *wh*-phrase is sub-extracted out of the subject constituent as acceptable in everyday Slovenian:

(1)  
*Čigavim predlogom se mu je [ugovarjati \_na oddelčnih sestankih] zdelo nesmisleno?*  
whose proposals cl.him is to-discuss at departmental meetings seemed pointless  
‘\*Whose proposals did to discuss at the departmental meetings seem pointless to him?’

The existing formal theories of syntactic locality (see, e.g. Chomsky 1986) generally predict that sub-extraction out of subject phrases, or subject islands, must be impossible, as illustrated by the English translation of (1). The acceptability of (1) and similar sentences suggests that some core structural factors that usually conspire to preclude these sentences fail to do so for some reason in Slovenian, and therefore raises a question as to why this might be so.

Taking Golden’s observation as a starting point, Stepanov et al. (to appear), in a larger-scale questionnaire study, investigated the pattern of grammaticality of similar subject island sentences using nominal subjects as in (2).

---

\* arthur.stepanov@ung.si

\*\* manca.msc@gmail.com

\*\*\* penka.stateva@ung.si

\*\*\*\* We thank the audience of the 8th Conference on Syntax, Phonology and Language Analysis (SinFonLJA 8) at the University of Ljubljana, and an anonymous Linguistica reviewer for helpful comments and suggestions. We also express our gratitude to everybody who contributed their native speaker judgments for this study. This work was partially supported by the EU 7th Framework project “Advancing the European Multilingual Experience” (ATHEME, Grant Agreement 613465), and by the Slovenian Research Agency program No. P6-0382.

- (2) a) *Kakšni je gospa mislila, da so [\_ študenti] prišli na žur?*  
 what-kind-of is lady thought that are students came on party  
 ‘\*What kind of did the lady think that students came to the party?’
- b) *Čigava je stric mislil, da je [\_ asistentka] ovirala policista?*  
 whose is uncle thought that is assistant hampered policeman  
 ‘\*Whose did the uncle think that assistant hampered the policeman?’

The results of that study revealed that sentences like (2) are perceived as substantially degraded by Slovenian speakers. On the basis of that, the authors concluded that at least in the nominal domain, Slovenian manifests a proper subject island effect. The results of that study, therefore, suggest that Slovenian is well-behaved with respect to the locality theories which predict degraded acceptability of subject island sentences.

At the same time, some of the results reported in Stepanov et al. (to appear) could in principle receive alternative interpretations because of two potentially relevant methodological issues. One issue regards the choice of the testing materials which involved Left Branch Extraction (LBE) of the nominal wh-specifier such as *kakšen* (“what kind of”) or *čigav* (“whose”). Since the seminal work of Ross (1967), it has been generally recognized that LBE is disallowed in languages like English (cf. *\*Whose did you see house?*). At the same time, LBE freely occurs in some Slavic languages including Russian and Bosnian/Croatian/Serbian. In Slovenian, examples like (3) are reported as acceptable in colloquial speech, and our informants agree on this and similar examples on which the materials for Stepanov et al. (to appear) were modeled:

- (3) a) *Čigave mi prinašaš [\_ pozdrave]?* Golden (1996: fn. 5)  
 whose me.dat. bring.2sg. greetings  
 ‘Whose greetings are you bringing me?’
- b) *Visoke je videl [\_ študente].* Bošković (2009: 70)  
 tall is seen students  
 ‘He saw tall students.’

(3) suggests that LBE is not precluded in Slovenian as well, and therefore, LBE should not incur additional penalty while testing extraction out of subject NPs, at least with respect to the wh-specifiers *kakšen* and *čigav*. At the same time, it is also the case that LBE appears to be generally more limited in Slovenian than in the above mentioned languages, although the literature remains somewhat vague as to the extent of this limitation. For instance, Bošković (2009: fn. 20) notes that factors including formal features of the element being extracted, agreeing auxiliary etc. affect the acceptability of examples like (3) in quite significant ways. Furthermore, Franks (2014) goes as far as to claim that Slovenian “generally eschews” LBE, referring to examples as in (4):

- (4) a) *\*Milojkina odhaja hči.*  
 Milojka’s leaves daughter  
 ‘Milojka’s daughter is leaving.’
- b) cf. *Milojkina hči odhaja.* Franks (2014: 162)

It might therefore be argued that the results of Stepanov et al. (to appear) concerning the degraded status of sentences involving wh-extraction out of subject islands may potentially be confounded with an additional constraint precluding LBE. The degraded status might then be due to either a) a subject island effect only, or b) LBE only; or c) a combination of subject island and LBE effects. This situation raises a further need to tease apart the subject island effect proper and a potentially intervening LBE effect. In order to do that, the situation with LBE in Slovenian must be further clarified.

The second potential issue is that the effect of subject islandhood reported in the study in Stepanov et al. (to appear) was found to be statistically marginal ( $p = 0.078$ ). The marginality of the subject island effect (reported  $p = 0.065$ ) was also a result of the large-scale study in Sprouse et al. (2015) which used materials involving sub-extraction out of subject NP in English and Italian. Both studies chose to interpret their result as significant. Nevertheless, we felt that statistical marginality of the results calls for a follow-up investigation replicating the subject island effect with a greater statistical power.

The present study has two main goals. Our first goal is to investigate whether Slovenian observes a subject island constraint, if the LBE factor is excluded. In essence, we wanted to replicate Stepanov et al.'s (to appear) findings concerning the islandhood of subject NPs, in principally non-LBE contexts. The second goal is to test whether Slovenian observes the constraint on LBE in NPs. Aside from the need to further clarify the relevant empirical pattern in this language, this goal is also pertinent in light of the recent proposals in the literature that postulate a one-way correlation between the absence of overt articles and allowing LBE in a language (Bošković 2005, 2008 and subsequent work). Slovenian does not have overt articles, and, therefore, presents an interesting case in terms of applicability of LBE, even though a priori it does not necessarily challenge the postulated correlation (see Section 4.2 for further discussion). Note that the two postulated goals are logically independent of each other in terms of contributing to the overall picture of locality in Slovenian. If it turns out that Slovenian does not allow LBE, **and** also observes the subject island constraint independently of the LBE factor, then we will have succeeded in ruling out the LBE factor from the inquiry into subject islands. If, however, Slovenian does not allow LBE, **and** it also does not observe the subject island constraint at the same time, that would mean that the original subject island results reported in Stepanov et al. (to appear) were actually due to LBE, not to the subject island constraint. Finally, if Slovenian comes out as allowing LBE, that would support the previous results concerning the subject islandhood status of Slovenian, at least in the domain of nominal subjects.

## 2. THE FACTORIAL DEFINITION OF ISLAND EFFECT

In order to determine whether Slovenian manifests a subject island effect, Stepanov et al. (to appear) used the factorial definition of island as developed in Sprouse et al. (2012, 2015). We employ the same definition of island in the present study as well. The rationale behind this definition lies in acknowledging that long-distance syntactic



dependencies in general, and island structures in particular, are syntactically complex structures whose comprehension may be affected, in particular, by the following two factors. Factor LENGTH reflects the length of the dependency between the fronted wh-phrase and its corresponding gap, that is, the position from which wh-movement takes place. This factor basically distinguishes extraction from a matrix clause and extraction from an embedded clause. The second factor, STRUCTURE, regulates whether a sentence contains an island sub-structure or not, independently of wh-movement. The factorial definition of island effect recognizes that each of these two factors negatively affects acceptability on their own. The character of this influence can be described in processing-related terms, on the assumption that the cognitive mechanisms engaged in processing island structures are subject to the same parsing constraints and limitations (e.g. related to working memory) that operate in any language under the strictly incremental character of syntactic processing. For instance, processing a longer dependency generally requires more memory resources than processing a shorter dependency. In a similar vein, processing a more complex structure such as a complex NP with rich internal structure, e.g. a relative clause, is generally more costly than processing a less complex NP such as *John*. Since each of the two factors has two values, the respective stimulus set involves four sentence types as shown below for the subject island (examples from Sprouse et al. 2015):

(5)

- a) *Who \_\_\_ thinks the speech interrupted the primetime TV show?*  
[non-island | matrix]
- b) *What do you think \_\_\_ interrupted the primetime TV show?*  
[non-island | embedded]
- c) *Who \_\_\_ thinks the speech about global warming interrupted the primetime TV show?*  
[island | matrix]
- d) *What do you think the speech about \_\_\_ interrupted the primetime TV show?*  
[island | embedded]

One line of thought in syntactic and processing literature maintains that island effects may be reduced to considerations of parsing efficiency alone (e.g. Hofmeister & Sag 2010, see also Hawkins 1999). Viewed in the context of the above two factors, this amounts to a cumulative, additive effect of those factors. That is, the degree of unacceptability that equals a sum of the degrees of unacceptability caused by each of these two factors alone (e.g. (5a-5d) = (5a-5b)+(5a-5c)), would imply that an island effect can be exhaustively modeled by these two factors. If the island effect is only due to these two factors, then a simple additive effect is all that is to be expected under these circumstances. If, however, the island effect exists over and above these processing considerations, then the factorial definition makes it possible to isolate it, in the form of a super-additive effect whereby the degree of unacceptability of an island construction such as (5d) is greater than the sum of the degrees of unacceptability caused by each of the two above mentioned factors alone, viz. (5a-5d) > (5a-5b)+(5a-5c). This superadditive effect

can be identified by using the measure known as *differences-in-differences* (DD), that is,  $DD = (5a-5d) - ((5a-5b)+(5a-5c))$ . If  $DD = 0$ , there is no island effect; if  $DD > 0$ , there is an isolated island effect independent of the above two factors (though the authors of the method do not provide a metric of islandhood based on the DD score). Alternatively, the superadditive effect can be detected as a statistically significant interaction of the LENGTH and STRUCTURE factors, in an ANOVA-type analysis. In general, a super-additive effect suggests that there exists a constraint over and above the processing-motivated LENGTH and STRUCTURE factors. An island effect can then be effectively teased apart from these processing-related factors. Thus the methodology based on the factorial definition of island offers a clear advantage over the standard methodology of data collection, which is not as sensitive to the potential influence of various factors in determining the grammaticality status of island sentences.

The prediction in our case is that, if the respective sentences in Slovenian manifest a true subject island effect, we expect  $DD > 0$  and a significant interaction between the two factors. If, in contrast, there is no true island effect, then it must be that  $DD = 0$  and there is no statistically significant interaction between the two independent factors.

### 3. THE PRESENT STUDY

As stated above, the present study has a dual purpose. On the one hand, we wanted to see whether Slovenian manifests the subject island effect if the LBE factor is excluded. Independently of that, we were also interested to determine if Slovenian manifests a LBE effect in interrogative and non-interrogative sentences. The two parts of the study are also slightly different methodologically. The first part is concerned with detecting an island effect as emerging from interaction of two independent factors, while the second is concerned with detecting an LBE effect as such. Consequently, we use the factorial definition of island effects for the first part of the study, but not for the second where we compare sentences with LBE with sentences without LBE. For the second part of the study we decided to ask a more comprehensive question, namely, whether the sensitivity of Slovenian speakers to the LBE contexts could possibly be non-trivially affected by a) the length of the respective dependency, contrasting matrix and embedded clauses, similarly to the subject island sub-experiment; and b) the type of LBE-triggering movement, contrasting *wh*-movement and non-*wh*-movement, the latter understood for the present purposes as displacement for reasons other than *wh*-movement. In essence, then, the present study comprises two sub-experiments in one.

#### 3.1 Materials

##### 3.1.1 Subject Islands

We used the factorial definition of island as a basis for tracking potential island effects in extraction out of subject NPs in Slovenian. This definition was implemented in a 2 x 2 design crossing factors LENGTH and STRUCTURE in the sense outlined above.

Stepanov et al. (to appear) tested for a subject island effect using materials involving LBE out of subject NPs, using a similar design (cf. (2)). In order to avoid potential LBE-related concerns, in this part of the study we used constructions involving PP extraction out of NPs modified by adjectives, as in the following example:

- (6) *Od koga si videl [veliko sliko \_] v dvorani?*  
of whom are.2sg seen large picture in hall  
‘Who did you see a large picture of in the hall?’

A sample set of examples is given in (7):

- (7)
- a) *Kdo \_\_ je mislil, da je [slika] visela v dvorani?* [NON-ISL. | MATR.]  
who is thought that is picture hung in hall  
‘Who thought that the picture hang in the hall?’
- b) *Kaj je Rok mislil, da je [\_] viselo v dvorani?* [NON-ISL. | EMB.]  
what is Rok thought that is hung in hall  
‘\*What did Rok think that hang in the hall?’
- c) *Kdo \_\_ je mislil, da je [velika slika Kosovela] visela v dvorani?* [ISL. | MATR.]  
who is thought that is large picture Kosovel hang in hall  
‘Who thought that a/the large picture of Kosovel hang in the hall?’
- d) *Od koga je Rok mislil, da je [velika slika \_\_] visela v dvorani?* [ISL. | EMB.]  
of whom is Rok thought that is large picture hang in hall  
‘\*Who did Rok think that a/the large picture of hang in the hall?’

The type of emdedded verb has been previously shown in the literature to affect the acceptability scores in subject island sentences (Polinsky et al. 2013). We controlled for potential variability in this domain by selecting only unaccusative and passive(-like) structures as embedded clauses in our Slovenian materials. In selecting unaccusative verbs, we followed a diagnostic suggested in (Marvin 2000): unaccusative verbs, unlike transitive ones can form past participle in this language. This is illustrated in (8) (Marvin’s examples (4b,c)):

- (8) a) *Videl sem žensko, prispelo danes zjutraj* [Past Part-unaccusative]  
seen am woman.acc. arrived today morning  
‘I saw a woman who arrived this morning.’
- b) *\*Videl sem žensko, napisalo knjigo* [Past Participle-transitive]  
seen am woman.acc. written book  
‘I saw a woman who wrote a book’

### 3.1.2 LBE: Materials

The LBE part of the study was implemented as a 2 x 2 x 2 design crossing the factors ‘LBE-hood’ (yes, no), LENGTH (matrix, embedded) and TYPE of movement (wh-,

non-wh). This resulted in eight conditions. The [wh-] subset of conditions is exemplified in (9), and the [non-wh-] subset is exemplified in (10):

- (9) a) *Kakšno žogo je Maja kupila \_\_, ko je šla v trgovino?* non-LBE / M  
 what-kind-of ball is Maja bought when is gone in market  
 ‘What kind of ball did Maja buy, when she went to the market?’
- b) *Kakšno žogo je gospod mislil, da je Maja kupila \_\_?* non-LBE / E  
 what-kind-of ball is man think that is Maja bought  
 ‘What kind of ball did the man think that Maja bought?’
- c) *Kakšno je Maja kupila \_\_ žogo, ko je šla v trgovino?* LBE / M  
 what-kind-of is Maja bought \_\_ ball when is gone in market  
 ‘What kind of ball did Maja buy, when she went to the market?’
- d) *Kakšno je gospod mislil, da je Maja kupila \_\_ žogo?* LBE / E  
 what-kind-of is man think that is Maja bought ball  
 ‘What kind of ball did the man think that Maja bought?’
- (10) a) *Rdečo kapo je Anka nosila \_\_, ko je spoznala Mateja.* on-LBE / M  
 red hat is Anka worn when is met Matej  
 ‘It is a red hat that Anka wore when she met Matej.’
- b) *Rdečo kapo je Matej mislil, da je Anka nosila \_\_.* non-LBE / E  
 red hat is Matej thought that is Anka wore  
 ‘It is a red hat that Matej thought that Anka wore.’
- c) *Rdečo je Anka nosila \_\_ kapo, ko je spoznala Mateja.* LBE / M  
 red is Anka worn hat when is met Matej  
 ‘It is a red hat that Anka wore when she met Matej.’
- d) *Rdečo je Matej mislil, da je Anka nosila \_\_ kapo.* LBE / E  
 red is Matej thought that is Anka wore hat  
 ‘It is a red hat that Matej thought that Anka wore.’

### 3.2 Questionnaires

We constructed eight sets of target sentences related to the subject islands (cf. (7)) using the same lexicalization for each set, eight sets of LBE-related sentences in the interrogative form, (cf. (9)) and eight sets of LBE-related sentences in the non-interrogative form (cf. (10)). Each series of eight sets was then distributed across eight lists using the Latin square procedure. The eight lists were then combined in pairs, which resulted in four master lists containing two sentence tokens (=lexicalizations) for each of the four conditions for each series, such that lexically related sentences never appeared in the same list. Thus each master list contained 24 target sentences that were not lexically related. Each master list was then supplemented with 24 filler sentences (half acceptable, half unacceptable, as judged by a linguist native speaker of Slovenian). This diversified the content of the questionnaires also minimizing possible rating biases. Two pseudo-random orders of each list were created, which resulted in 8 unique questionnaires of 48 items. Additionally, it was ensured that the first four items in each list are fillers.

### 3.3 An Acceptability Rating Task

In this study, we used the task of magnitude estimation. Magnitude estimation is a method of subjective evaluation whereby a participant evaluates some gradable property (e.g. intensity of light) relative to some available standard, by assigning a numerical value on the basis of a subjective judgment, in relation to the numerical value assigned to the standard (Stevens 1975). The subjects are not restricted either in the range of numerical values that they are allowed to give (on the positive number scale), or in the granularity of the numerical scale adopted by each participant for the purposes of the experiment. We used a version of the magnitude estimation task adapted for judging acceptability of sentences (Bard et al. 1996). This task is well suited to the present study because of its capability to capture a potentially greater variability and range of acceptability ratings by using the unbounded positive number line.

The task began with a training session, the goal of which was to familiarize the participants with the concept of magnitude estimation. During the training session, the subjects were offered to estimate the length of seven straight lines relative to the given line to which the numerical score 100 was assigned. Subjects were instructed to rely only on their subjective intuitions in evaluating the length of the lines; if the line seemed, for instance, twice as large as the standard, they were encouraged to give a value 200, and if it seemed about one third as large, then they would give a value 30. Both whole and decimal numbers could be used.

The training was followed by a sentence-rating questionnaire comprising the materials as described above. In our study, participants were presented with a reference sentence and a numeric value representing its acceptability. In our case, the sentence was (11), and it was pre-assigned the value 100 (note that the number itself does not imply any particular acceptability status; this point was also stressed in the instructions).

- (11) *Proti kateremu pravilu je Klara mimogrede rekla, da je Cene protestiral?*  
against which rule is Klara in-passing said that is Cene protested  
'Against which rule did Klara say in passing that Cene protested?'

The participants were then instructed to indicate the acceptability of each of the subsequent sentences relative to the score assigned to the standard. The participants were also instructed to judge the sentences following their first intuitive hunch, not the normative standards for Slovenian, and not to dwell on particular sentences as they go along.

The study was conducted in the form of a paper survey. The reference remained visible throughout the entire procedure by being placed on top of each page of the questionnaire and separated by a line from the rest of the stimuli. Participants were under no time constraints to complete the task. On average, the surveys were completed within 25 minutes.

### 3.4 Participants

Forty adult native speakers (thirty-two females) of Slovenian aged 19-53 (mean age: 28.75) participated in the experiment voluntarily and anonymously. All participants

had normal or corrected to normal vision. They were naïve to the purposes of the study. None of the participants had previously taken part in similar experiments for at least two years. The participants completed the task individually under the experimenter’s supervision. No participant data were excluded from the analysis.

### 3.5 Statistical Procedures

Prior to analysis, the raw numerical ratings from each participant were z-score transformed. The z-score transformation converts each participant’s ratings to a standardized score, in which each transformed rating represents the number of standard deviations by which the corresponding raw rating is different from that participant’s mean rating. This kind of conversion eliminates potential scale biases between participants (such as choosing different ranges of values among participants or using one end of the scale), and therefore allows for a cleaner comparison of the participants’ performance.

For the statistical analyses, we used linear mixed-effects models (Baayen et al. 2008). LENGTH and STRUCTURE were used as fixed factors for the subject island part of the study, and “LBE-hood”, LENGTH and movement TYPE were used as fixed factors for the LBE part of the study. In both parts of the study, participants and items were entered as random factors into the models. We report  $p$  values based on the likelihood-ratio test whereby a model containing the fixed effect of interest is compared to a model that is identical in all respects except the fixed effect in question. Analyses were performed using the “lme4” package (Bates et al. 2014) in R (R Core Team 2014).

For the subject island portion of the experiment, we also computed DD scores for each participant, on the basis of which we calculated mean DD scores for each island as a non-standardized effect-size measure for the island types under question (see Section 2).

## 3.6 Results

### 3.6.1 Subject Islands

Linear mixed-effects modeling revealed a main effect of factor LENGTH, as well as a main effect of factor STRUCTURE. Unsurprisingly, these two factors were found to play a role in assessing the grammaticality of the island sentences. Furthermore, we found that these two factors significantly interact with each other in a superadditive manner. Under the factorial definition of island, the presence of a robust and clear superadditive effect that obtains over and above the influence of each of these two factors alone suggests a true island effect independent of the processing considerations as well as free from a potential confound in the form of the LBE effect. In addition, we found that  $DD > 0$ , consistently with the superadditive character of this effect.

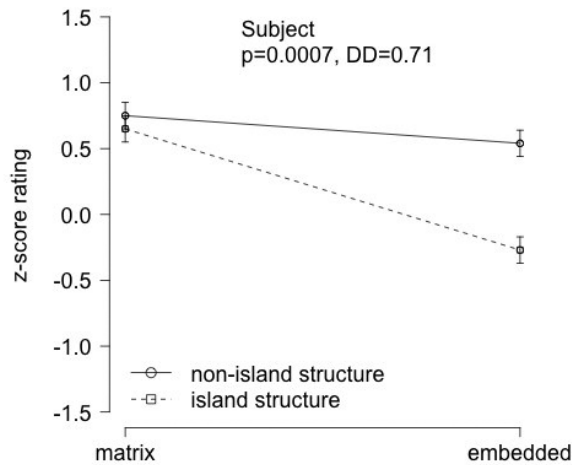
We also estimated the processing costs of LENGTH and STRUCTURE separately by computing the relevant pairwise comparisons: the length cost was identified with a pairwise comparison of NON-ISLAND | MATRIX and NON-ISLAND | EMBEDDED conditions, and

the structure cost was identified with a pairwise comparison of ISLAND | MATRIX and NON-ISLAND | MATRIX conditions (see also Sprouse et al. 2012). The cost effects of LENGTH and STRUCTURE came out not significant for the subject island structure. The results of this part of the study are summarized in Table 1 and Figure 1.

Table 1:  $\chi^2$ ,  $t$  and  $p$ -values for the linear mixed-effects models fitting the subject island data

<i>FULL 2 X 2 MODEL</i>	SUBJECT	
	$\chi^2$	$p$
main effect of LENGTH	17.522	<0.0001
main effect of STRUCTURE	12.213	0.0005
interaction LENGTH x STRUCTURE	<b>11.352</b>	<b>0.0007</b>
<i>PAIRWISE COMPARISONS</i>	$t$	
LENGTH (STRUCTURE=non-island)	1.390	0.5138
STRUCTURE (GAP=matrix)	0.614	0.9268

Figure 1: Interaction plot for the subject island



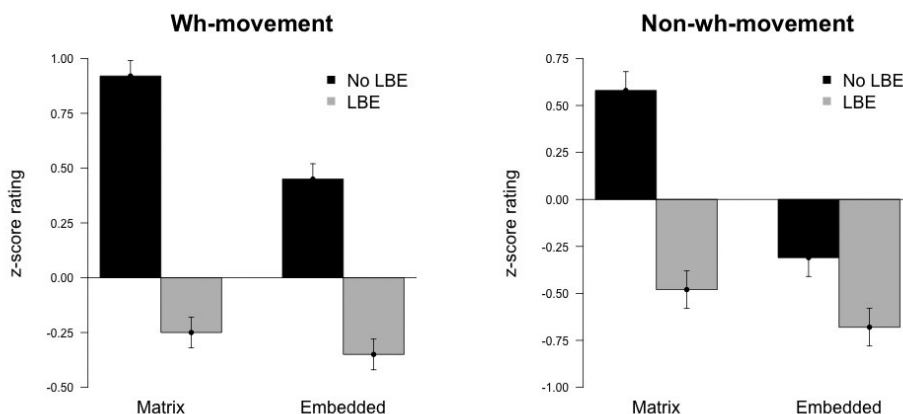
### 3.6.2 LBE

We first evaluated the overall 2 x 2 x 2 model for the LBE sub-experiment. Main effects were observed for each of the three factors involved, that is, LBE-hood, LENGTH and TYPE. In other words, each of the above factors emerged as a significant predictor of the acceptability scores. There was also a significant three-way interaction among these factors. The results are summarized in Table 2 and Figure 2.

Table 2:  $\chi^2$ ,  $t$  and  $p$ -values for the linear mixed-effects models fitting the wh-LBE data

<i>FULL 2 X 2 X 2 MODEL</i>	$\chi^2$	$p$
main effect of LBE-hood	75.797	<0.0001
main effect of LENGTH	27.848	<0.0001
main effect of TYPE	27.418	<0.0001
interaction LBE-hood x LENGTH x TYPE	28.914	<0.0001

Figure 2: Z-score comparison of the sentences involving and not involving LBE.



To better understand this three-way interaction, we then considered two smaller 2 x 2 models crossing factors LBE-hood and LENGTH and pertaining to wh-movement and non-wh-movement, respectively. We found a robust main effect of LBE-hood in both interrogative and non-interrogative sentences. There was also a main effect of LENGTH in both construction types, with extraction out of matrix clauses receiving higher score than out of embedded clauses. Planned pairwise comparisons confirmed that LBE constructions were judged significantly lower than non-LBE sentences, in both matrix and embedded contexts, and for both movement types ( $p < 0.004$  for all pairs). Furthermore, we observed a significant interaction between LBE-hood and LENGTH, suggesting that the length of a dependency affects acceptability of the LBE structures.

We also constructed two 2 x 2 models crossing factors LBE-hood and movement TYPE, pertaining to matrix and embedded clauses, respectively. LBE-hood again had a main effect, and so did TYPE, for each of the clausal types. In other words, it matters for the participants whether extraction takes places in the form of wh-movement or another movement type, irrespective of LBE. Interestingly, however, a significant interaction between factors LBE-hood and movement type was observed only for the embedded, though not for the matrix, clauses, suggesting that the type of movement



affects acceptability of LBE structures only in the former. Post-hoc pairwise comparisons also showed that LBE structures were given scores significantly lower in the non-wh-movement contexts compared to the wh-movement contexts, as far as embedded clauses ( $p < 0.02$ ), but not matrix clauses ( $p = 0.31$ ), are concerned. These four 2 x 2 models are summarized in Tables 3 and 4.

Table 3:  $\chi^2$  and  $p$ -values for the 2 x 2 models crossing LBE-hood and LENGTH

2 X 2 MODELS	WH-EXTRACTION		NON-WH-EXTRACTION	
	$\chi^2$	$p$	$\chi^2$	$p$
main effect of LBE-hood	60.063	<0.0001	26.657	<0.0001
main effect of LENGTH	12.396	0.0004	18.123	<0.0001
interaction LBE-hood x LENGTH	7.0925	0.0077	12.479	0.0004

Table 4:  $\chi^2$  and  $p$ -values for the 2 x 2 models crossing LBE-hood and TYPE

2 X 2 MODELS	MATRIX		EMBEDDED	
	$\chi^2$	$p$	$\chi^2$	$p$
main effect of LBE-hood	57.826	<0.0001	31.026	<0.0001
main effect of TYPE	9.0547	0.00262	28.323	<0.0001
interaction LBE-hood x TYPE	0.3736	0.5411	9.0822	0.0026

### 3.7 Discussion

#### 3.7.1 Subject Islands

Our goal in this part of the study was to test for subject island effects in Slovenian excluding the LBE factor. There are two main results of the sub-experiment involving subject islands. First, we establish that there is a robust subject island effect in Slovenian. This effect shows up in the form of a superadditive effect as a result of the interaction of the independent factors LENGTH and STRUCTURE. This result largely replicates the results reported in Stepanov et al. (to appear), with two important differences, each of which relates to the respective concern posed in the beginning of this study. First, the latter work used materials that involve extraction out of subject NP in the form of LBE (cf. (2)). Our present concern was that the lowered acceptability on the subject island sentences reported in that study could in principle be interpreted in at least three ways: a) being due to a combination of LBE and subject island; b) due to LBE alone; and c) due to a subject island alone (see Section 1). The present study teases apart these possibilities. Since there is no LBE involved in our materials, a potential LBE confound is therefore eliminated, and the observed effect can reasonably be attributed to the subject island alone.

The second difference is that the present study reports a cleaner and more robust effect pertaining to the subject island than that reported in Stepanov et al. (to appear). In the latter study, in which extraction out of subject NP was an instance of LBE, the observed superadditive effect was of marginal significance. In the present study involving PP extraction out of NP, the effect is shown to be statistically significant, eliminating potential ambiguities concerning its interpretation.

The contributing costs of factors LENGTH and STRUCTURE into the overall unacceptability of respective sentences, as estimated by pairwise comparison tests, were found insignificant in the present study (see Table 1), replicating the findings of Stepanov et al. (to appear). This state of affairs suggests that the observed true island effect in subject islands in Slovenian is due to reasons beyond these processing factors, namely those that have to do with the grammar proper rather than performance. Earlier, using similar materials, Sprouse et al. (2012) reported that the factor STRUCTURE did not incur an independent processing cost in subject island sentences in English. We speculate that an explanation of this state of affairs might lie in the constructed syntactic complexity, and the related processing complexity, of the relevant noun phrases. The structures used in Sprouse et al. (2012) to represent the NON-ISLAND and ISLAND values of the factor STRUCTURE had a shape such as *what* vs. *the speech about global warming*, respectively (see (5)). Thus the added complexity in the ISLAND condition comes from the prepositional phrase (PP) *about global warming*. In the Slovenian materials used the present study, the contrast between the two conditions lies in the presence of an adjective, e.g. *velika slika* "picture" vs. *velika slika Kosovela* "a large picture of Kosovel" (cf. (7)). Possibly, this added syntactic complexity is not sufficient to incur a significant processing cost either in English or in Slovenian. This is different, for instance, from *wh*-islands, which typically represent a clausal piece of structure, hence, presumably, are a priori more syntactically complex (see the above studies for more details).

Our results also indicate that the factor LENGTH does not incur an independent processing cost. In a similarly constructed study of Sprouse et al. (2012) with English materials, LENGTH was found to incur such independent cost in the subject island-related sentences. Stepanov et al. (to appear) speculated that a relevant cross-linguistic difference might lie in the nature of the testing materials. The difference between our materials in the present study and those used in the English study is that our materials involve D-linked *wh*-phrases, that is, (the Slovenian counterpart of) *which*-phrases, whereas in the reported English study bare *wh*-words such as *what* and *who* are extracted (cf. (5) vs. (7)). D-linked phrases are generally known to be subject to more liberal constraints on extraction than bare *wh*-words. Processing-wise, it has been demonstrated that items that are richer in featural composition leave a longer and more robust trace in the working memory, and consequently are subject to a slower memory decay compared to items that have less relevant features (e.g. Hofmeister and Vasishth 2014). Thus a D-linked phrase having a richer featural make-up may be able to linger in the memory for a longer time, overcoming potential effects of dependency length. The observed lack of independent processing cost incurred by LENGTH could possibly be

due to that. However, this does not mean that indefinitely increasing the length of the dependency will have no effect on the acceptability whatsoever: there must be some threshold value that even the D-linked character of the *wh*-phrase cannot overcome. This is suggested, in particular, by our results concerning LBE effects below. It should also be mentioned that, in the study of Stepanov et al. (to appear), *wh*-islands in Slovenian were found to properly incur independent processing costs of both LENGTH and STRUCTURE, as expected under this kind of considerations.

### 3.7.2 LBE

The results of the second part of our study strongly suggest that Slovenian observes a constraint on LBE, in interrogative as well as non-interrogative sentences. Speakers generally dislike extraction of a *wh*- as well as a non-*wh*-specifier out of NP in the object position. Furthermore, factor LENGTH plays a role as well: sentences with matrix LBE are judged more acceptable than sentences with embedded LBE, in both *wh*- and non-*wh*-versions. This is different from the subject island case where LENGTH was not found to be a significant factor. It should be noted, however, that, all else equal, the dependencies in the subject island-related sentences are a priori shorter than those in our LBE-related sentences (both involving and not involving LBE) where extraction out of the object position takes place. This is because extraction from subject NPs, in a canonical SVO configuration, a priori incurs a shorter dependency than extraction from object NPs. Therefore, a LENGTH effect observed in the LBE-related sentences is not surprising, and is on a par with a similar effect involving extraction from object in various types of islands, e.g. *wh*-island or complex NP island (see Sprouse et al. 2012, 2015; Stepanov et al. to appear for discussion).

Our results also suggest that the non-*wh* extraction sentences are perceived by the speakers as significantly worse than the *wh*-extraction sentences (see Section 3.6.2). In other words, the effect of movement type suggests that non-*wh*-fronting is generally disliked by the speakers regardless of LBE. This might reflect a genuine grammatical and/or processing constraint distinguishing among these movement types. As further elucidation of this putative constraint requires a more fine-grained excursus into theoretical details concerning the clausal and information structure of Slovenian, we leave it for future research, noting its potential importance in the context of computational mechanisms and triggers for various types of syntactic movement and their manifestation in this language. An alternative possibility is that this result might be due to the presentation format of our study. It is well known that fronted non-*wh*-constituents in Slavic languages usually bear an additional informational burden (e.g. contrastive focus) that can be properly construed only if an appropriate discourse context is provided. Since the relevant sentences were presented for evaluation to our participants context-free, it is possible that the participants gave such sentences a low score because of the lack of such context and the ensuing difficulty to assign these sentences a proper syntactic and semantic analysis (in contrast, sentences with *wh*-movement do not require such articulated context). Therefore, a follow up study regarding the influence of the movement type on the LBE structures might be in

order, perhaps using a different experimental methodology that would test this and other alternative interpretations.

## **4. A WORD ON POTENTIAL THEORETICAL CONSEQUENCES**

### **4.1 Subject Islands**

In contrast to the earlier claims (see Section 1), Slovenian appears to be well-behaved with respect to the subject island constraint. With respect to subject NPs, Slovenian can be placed on a par with many other languages manifesting the same constraint on extraction. This result extends so far to nominal subjects only. We have not tested sentential subjects in our study and it remains to be seen whether our conclusion can also be extended to those.

It should be noted that Golden (1995, 1996, 1997) reports examples of wh-extraction out of adjuncts in Slovenian as ungrammatical, and this comports well with similar observations from other languages in the literature. The fact that Slovenian manifests a subject island effect, taken together with the reported degradation of adjunct island sentences in the earlier literature, suggests that Slovenian is well-behaved with respect to the Condition on Extraction Domain (CED) in its original formulation (cf. Huang 1982). The CED predicts that subjects and adjuncts are a natural class of domains immune to sub-extraction from them. Some languages have since been shown in the literature to manifest a diverging behavior with respect to acceptability of sentences involving sub-extraction from subjects, whereas languages tend to be uniform in their ban on extraction from structural adjuncts (see e.g. Stepanov 2007). Each case of such divergence therefore presents an a priori challenge and an interesting empirical test case to the CED as a principle of grammar. If the earlier claims to the effect that Slovenian does not observe a subject island constraint were confirmed, Slovenian would then present another be an interesting case to study with respect to the nature of the divergence. The present study demonstrated, however, that Slovenian presents no such challenge to the CED from the part of nominal subject island constructions.

### **4.2 LBE**

A number of authors argued that the possibility for LBE correlates with the lack of articles in a given language (see, e.g. Uriagereka 1988; Corver 1992; Bošković 2005). For instance, the Germanic languages generally have articles, and do not allow LBE. In contrast, the Slavic languages such as Serbian or Russian do not have articles, and they permit LBE. Now, Slovenian is a language that does not have overt articles of the kind found in Germanic languages. From this perspective, Slovenian might appear problematic in light of the observed constraint on LBE, which makes it more similar to the Germanic languages. However, Bošković (2005, 2008) argues that the correlation is one-way only: an articleless language may, but does not have to allow, LBE. Japanese is an example of the latter. In other words, the set of articleless languages is a superset of languages that allow LBE. Bošković (2005, 2008) also argues that languages that do

not have articles actually lack the DP layer of the structure of nominal phrases. Thus languages with articles have DPs, while languages without articles have NPs only. In Bošković's system, then, the lack of the DP layer is a necessary (but not sufficient) condition for LBE.<sup>1</sup>

It is beyond the scope of the present paper to go into the details of the productive NP/DP debate in the literature (see also fn. 1). If the main proposal is on the right track, then we face two potential theoretical possibilities. One is that Slovenian is like Japanese, meaning that whatever principle (other than the parametric variation in the NP/DP structure) accounts for the LBE-hood and the absence of the articles, it is irrelevant in the case of Slovenian. Another possibility, still within the NP/DP paradigm, is diachronic: Slovenian may be a language that is about to change from an "NP language" into a "DP language", one that may eventually develop a full-fledged article system. Bošković (2008: fn. 23) acknowledges this possibility. Some independent phenomena from the Slovenian syntax also indirectly suggest that this possibility is viable (Marušič and Žaucer 2014). The diachronic grammar-in-the-flux possibility could, in principle, also account for the apparently inconsistent character of the informally reported judgments, whereby certain instances of LBE are allowed (cf. (3)) whereas others are not. These possibilities will need to be distinguished in light of additional evidence that should emerge in future investigations.

## 5. CONCLUSION

In this study we aimed at strengthening the empirical base for the theories of syntactic locality by investigating two domains of syntactic locality in Slovenian. First, building on the previous findings concerning the presence of the subject island effect in Slovenian in Stepanov et al. (to appear), we replicated these findings while also excluding the LBE factor that was a potential confound in the previous study. We used a different set of sentences not involving LBE and found that the subject island effect in Slovenian persists, even in a more robust manner than what was observed before. Second, independently of that, we also asked whether Slovenian observes a constraint on LBE by using materials involving *wh*- and non-*wh*-displacement from the object position, both in matrix and embedded environments. We found that Slovenian speakers are sensitive to the constraint on LBE, modulated also by the length of the respective dependency (factor LENGTH), and that this sensitivity persists across *wh*- as well as non-*wh*-dependencies.

These results contribute to the growing body of evidence concerning syntactic locality domains in Slovenian. As noted in the beginning of this article, there are reasons

---

1 More recent accounts of the phenomenon hold that the presence of a phase (not necessarily a DP phase) above the respective NP may block LBE in a language that otherwise allows it (see e.g. Bošković 2014 for relevant evidence from Serbian/Croatian/Bosnian; see Chomsky 2001 and later works on the concept of phase). We agree with an anonymous reviewer that this type of account offers a potentially promising venue for analyzing the seemingly diverging data patterns concerning LBE in Slovenian.

to believe that Slovenian observes constraints on certain syntactic islands including e.g. adjunct island, complex NP island or coordinate structure constraint. On the other hand, another recent finding reported in Stepanov et al. (to appear) was that Slovenian speakers do not observe the wh-island constraint. Considered together with the results of the present study, the wh-island emerges as the only island type in the Slovenian grammar whose status deviates from the expected range. This suggests one potential focus and provides a good continuation point for further theoretical studies of locality in Slovenian. Another interesting domain concerns further investigation of LBE involving extraction of different types of adjectival and/or adverbial specifiers.

## References

- BAAAYEN, R. Harald/Douglas J. DAVIDSON/Douglas M. BATES (2008) “Mixed-effects modeling with crossed random effects for subjects and items.” *Journal of memory and language* 59, 390–412.
- BARD, Ellen Gurman, Dan ROBERTSON, and Antonella SORACE (1996) “Magnitude estimation of linguistic acceptability.” *Language* 72, 32–68.
- BATES, Douglas/Martin MAECHLER/Ben BOLKER/Steve WALKER (2014). lme4: Linear mixed-effects models using Eigen and S4\_. R package version 1.1-7.
- BOŠKOVIĆ, Željko (2005) “On the locality of left branch extraction and the structure of NP.” *Studia Linguistica* 59, 1–45.
- BOŠKOVIĆ, Željko (2008) “What will you have, DP or NP?” *Proceedings of the North East Linguistic Society (NELS)* 37, 101–114.
- BOŠKOVIĆ, Željko (2009) “The NP/DP analysis and Slovenian.” In: *Proceedings of the University of Novi Sad Workshop on Generative Syntax*, 53–73.
- BOŠKOVIĆ, Željko (2014) “Phases beyond clauses.” In: L. Schurcks/A. Giannakidou/U. Etxeberria (eds), *The nominal structure in Slavic and beyond*. Berlin: De Gruyter, 75–128.
- CHOMSKY, Noam (1986) *Barriers*. Cambridge, MA: MIT press.
- CHOMSKY, Noam (2001) “Derivation by phase.” In: Michael Kenstowicz (ed.), *Ken Hale: a life in language*. Cambridge, MA: MIT Press, 1–52.
- CORVER, Norbert. (1992) “Left branch extraction.” *Proceedings of NELS* 22, 67–84.
- FRANKS, Steven (2014) “The Slovenian orphan accusative, empty pronouns and noun phrase structure.” In: L. Schurcks/A. Giannakidou/U. Etxeberria (eds), *The nominal structure in Slavic and beyond*. Berlin: De Gruyter, 129–182.
- GOLDEN, Marija (1995) “Interrogative wh-movement in Slovene and English.” *Acta Analytica* 14, 145–187.
- GOLDEN, Marija (1996) “K-premik in skladenjski otoki v slovenski skladnji.” *Razprave SAZU, Razred II*, 15, Ljubljana: SAZU, 237–253.
- GOLDEN, Marija. (1997a) *O jeziku in jezikoslovju*. Ljubljana: Filozofska fakulteta.
- GOLDEN, Marija. (1997b) “Multiple wh-questions in Slovene.” In: *Proceedings of Formal Approaches to Slavic Linguistics 4 (The Cornell meeting)*. Ann Arbor, MI: Michigan Slavic Publications, 240–266.



- HAWKINS, J. A. (1999) “Processing complexity and filler-gap dependencies across grammars.” *Language* 75, 244–285.
- HOFMEISTER, PHILIP/Ivan A. SAG. (2010). “Cognitive constraints and island effects.” *Language* 86, 366–415.
- HOFMEISTER, Philip, and Shravan VASISHTH. 2014. “Distinctiveness and encoding effects in online sentence comprehension.” *Frontiers in Psychology: Language Sciences* 5. doi:10.3389/fpsyg.2014.01237.
- HUANG, C.-T. James (1982) *Logical relations in Chinese and the theory of grammar*. Doctoral dissertation, Cambridge, Mass.: MIT.
- MARVIN, Tatjana (2000) “Past participles in reduced relatives.” In: *Proceedings of ConSOLE IX*. Lund: University of Lund, 141–156.
- MARUŠIČ, Franc/Rok ŽAUCER (2014) “A definite article in the AP - evidence from colloquial Slovenian.” In: L. Schurcks/A. Giannakidou/U. Etxeberria (eds), *The nominal structure in Slavic and beyond*. Berlin: De Gruyter, 183–208.
- POLINSKY, Maria/Carlos G. GALLO/Peter GRAFF/Ekaterina KRAVTCHENKO/Adam Milton MORGAN/Anne STURGEON (2013) “Subject islands are different.” In: J. Sprouse (ed.), *Experimental syntax and island effects*. Cambridge: Cambridge University Press, 286–309.
- R CORE TEAM. (2014). *R: A Language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. <http://www.R-project.org/>
- ROSS, John Robert (1967) *Constraints on variables in syntax*. Doctoral dissertation. Cambridge, MA: MIT.
- SPROUSE, Jon/Matt WAGERS/Colin PHILLIPS (2012) “A test of the relation between working-memory capacity and syntactic island effects.” *Language* 88, 82–123. doi:10.1353/lan.2012.0004.
- SPROUSE, Jon/Ivano CAPONIGRO/Ciro GRECO/Carlo CECCHETTO (2015) “Experimental syntax and the variation of island effects in English and Italian.” *Natural Language & Linguistic Theory* 34/1, 1–38. doi:10.1007/s11049-015-9286-8.
- STEPANOV, Arthur (2007) “The end of CED? Minimalism and extraction domains.” *Syntax* 10, 80–126.
- STEPANOV, Arthur/MUŠIČ, Manca/Penka STATEVA (to appear). “Two (non-) islands in Slovenian: A study in experimental syntax.” *Linguistics*.
- URIAGEREKA, Juan. (1988). *On government*. Doctoral dissertation. University of Connecticut.

### Summary

#### ASYMMETRIES IN SUB-EXTRACTION OUT OF NP IN SLOVENIAN: A MAGNITUDE ESTIMATION STUDY

In this work, we aim to clarify the empirical paradigm that bears on two aspects of syntactic locality in Slovenian. First, building on previous work, we investigate how robustly Slovenian observes the syntactic locality constraint precluding constituent

sub-extraction out of subject noun phrases. Second, we ask whether Slovenian allows Left Branch Extraction in interrogative and non-interrogative sentences. To elucidate both issues, we conducted a magnitude estimation study, the results of which support our previous claim that there is a subject island effect in Slovenian. Furthermore, our results suggest that Slovenian disallows Left Branch Extraction, in contrast with some other Slavic languages. We also discuss theoretical consequences of our empirical findings.

**Keywords:** syntactic island, Left Branch extraction, magnitude estimation, Slovenian

#### Povzetek

### NESOMERNOSTI PRI PREMIKU IZ SLOVENSKE SAMOSTALNIŠKE ZVEZE: ŠTUDIJA PO METODI OCENE MAGNITUDE

V pričujočem članku poskušamo razjasniti empirično paradigmo, ki vpliva na dva vidika skladske lokalnosti v slovenščini. Prvič, izhajajoč iz predhodne raziskave, smo želeli podrobneje proučiti, kako močno slovenščina upošteva omejitev skladske lokalnosti, ki preprečuje premik iz osebkovih samostalniških zvez. Drugič, ugotoviti smo želeli, ali slovenščina dopušča premik pridevnika iz samostalniške zveze v vprašalnih in nevprašalnih povedih. Da bi odgovorili na obe zastavljeni vprašanji, smo izvedli študijo po metodi ocene magnitude. Rezultati študije potrjujejo naše prejšnje zaključke, da slovenščina izkazuje prepoved premika iz osebkove zveze. Naši rezultati tudi pokažejo, da slovenščina ne dovoljuje premika pridevnika iz samostalniške zveze in tako kaže razločke z nekaterimi drugimi slovanskimi jeziki. Na koncu članek obravnava izsledke raziskave z vidika razvoja teorije skladnje.

**Ključne besede:** skladski otoki, premik pridevnika iz samostalniške zveze, metoda ocene magnitude, slovenščina







## LOW VOWEL “LENGTHENING” IN HUNGARIAN\*\*

### 1. INTRODUCTION

The alternation in Hungarian in which stem-final short low vowels ([ɔ] and [ɛ]) alternate with their long counterparts ([a:] and [e:], respectively) in suffixed forms of the stem, is usually referred to as Low Vowel Lengthening (LVL). LVL is a productive alternation, insensitive to vowel harmonic properties of the stem. It is illustrated in *Table 1*.

Table 1: Low Vowel Lengthening in Hungarian

Nominative Singular	Accusative Singular	Meaning
fɔ	fa:t	'tree'
kuɔ	kuca:t	'dog'
mɔʃku	mɔʃka:t	'cat'
pejvɔ	pejva:t	'chaff'
kɔlodɔ	kɔloda:t	'stocks (form of punishment)'
kefe	kefe:t	'brushN'
piske	piske:t	'gooseberry'
ʃerte	ʃerte:t	'bristle'
ogre	ogre:t	'ogre'
remete	remete:t	'hermit'

There have been several attempts at explaining this phenomenon, but certain obstacles have not been successfully overcome, even though it has been approached differently by different authors. Also, it has been analyzed as both shortening (Abondolo 1988, Rebrus 2000) and lengthening (Siptár and Törkenczy 2000) in the literature, but no general discussion has been presented on the differences between these two approaches. In the following, the most problematic cases of LVL will be identified, which will be followed by a comparison of theoretically possible groups of analyses. Finally, a new approach will be introduced.

\* ildi.szabo@nyu.edu

\*\* The work in this paper was presented and written while I was attending Eötvös Loránd University, Budapest, and it is part of the research conducted for the Hungarian National Scholarly Circle. Therefore I would like to thank – besides my anonymous reviewers – my supervisors: Péter Rebrus and Miklós Törkenczy. All mistakes are mine.

## 2. DIFFICULTIES

The following observation can be made about Hungarian phonotactics:

- (1) There are no words in Hungarian that end in [a:] or [e:].

There are two aspects of it which proved to be problematic, yet crucial for accounts of LVL. First, Novák (1999) provides a functional motivation for LVL. In Hungarian analytical and “quasi-analytical” suffixes can appear with linking vowels. These vowels are usually mid ones, yet a closed class of consonant-final stems (called Lowering Stems) take low linking vowels (e.g. *gáz-ok* ‘gas-PL NOM’ vs. *ház-ak* ‘ház-PL NOM’, *gáz* is a regular stem, *ház* is a LS in Hungarian). LVL is needed to ensure the reconstructability of the stem and the identification of stem-morpheme boundary. Therefore Lowering Stems should be easily fit into any analysis given for LVL.

Another peculiarity for which any approach of LVL should account for is the case of stems that do not trigger LVL, but still harmonize with the stem.

Table 2: Not triggering LVL, but harmonizing with the stem

kucɔʃa:g	*kucɔʃe:g	<i>kutyaság</i> ‘dogness’
mɛdvɛsɛ:g	*mɛdvɛsa:g	<i>medveség</i> ‘bearness’
lilɔʃa:g	*lilɔʃe:g	<i>lilaság</i> ‘purpleness’
fɛkɛtɛsɛ:g	*fɛkɛtɛsa:g	<i>feketeség</i> ‘blackness’

In such cases, separate phonological domains are often proposed for stem and suffix – in order to explain the absence of LVL – but then the application of Vowel Harmony is unexpected. Separately, these problems have been previously addressed in the literature, but no analysis so far proposed in the literature gave a solution for both of them.

## 3. SHORTENING, LENGTHENING OR PHONEME C?

The question of whether a shortening or a lengthening approach is more preferable in general, has so far been devoted little attention in the literature. Three options seems possible here: deriving alternating forms A and B from an underlying A, an underlying B or from a third, abstract phoneme (C). The third option has to be instantly rejected here, because in this case the data do not justify or necessitate such an approach and applying such a solution would add to the complexity and abstractness of the system and raise the old problem of generative phonology being unnecessarily abstract (Kiparsky 1968).

Positing /a:/ and /ɛ:/ as underlying and deriving the [ɔ]’s and [ɛ]’s of the Nominative forms from them by rule would mean analysing LVL as a shortening process in

which the underlying /a:/ or /e:/ is shortened word finally.<sup>1</sup> An analysis like this would suppose a phonologically less marked sound to be found underlyingly. This in line with cross-linguistics patterns in the case of the [ɔ] ~ [a:] alternation, since /a:/ undoubtedly appears in far more languages than /ɔ/ does. It would also entail that the base form appears in more marked forms and the less marked one is the one that is derived which would be somewhat unfavorable.

Such an analysis would not need a phonotactic constraint to arrive at a situation described in (1), which would seem tempting. Word finally, long and short low vowels are neutralized in this case. This, however, if more closely examined, is somewhat puzzling. The [ɔ] vs. [a:] and [ɛ] vs. [e:] contrast is a heavily loaded and also very salient contrast, and its neutralization, in lieu of a (functionally motivated, but otherwise arbitrary) phonotactic constraint, has to be accounted for by other arbitrary means. In this case there should also be a group of stems that, in the Nominative Singular, end in a short low vowel that systematically fails to lengthen in suffixed forms. These would be the stems that end in a short low vowel underlying and would question the efficiency of LVL for stem identification. Therefore, even if this move is tempting at first glance, there is no real structural advantage of not proposing a phonotactic constraint like (1) in shortening approaches.

The traditional approach to Low Vowel Lengthening is – as the name also shows – analyzing it as lengthening process which derives the [a:] and [e:] from the /ɔ/ and /ɛ/, respectively. Approaches of this kind, by positing /ɔ/ and /ɛ/ as underlying, posit the more cross-linguistically marked sounds to be underlying but the underlying form surfaces in less marked forms of the paradigm, for instance, in the Nominative Singular that has no overt marking in Hungarian. A lengthening analysis would also necessarily assume that long low vowels are banned word finally in Hungarian by a phonotactic constraint. Therefore the constraint stipulated in (1) has to be perceived as a static phonotactic rule rather than a dynamic one; the lack of long low vowels in a word final position is not a result of a neutralization process, but the possibility of their presence is excluded altogether.

Lengthening approaches have one strong advantage over shortening ones. While shortening approaches share the dynamics of the stem having the possibility of being long in itself, but word final position blocks this process or actively reduces the vowel, these analyses by default assume that having an overt suffix is a protection from this effect. If there is a suffix that does not trigger LVL, it has to be motivated through the same condition or mechanism and therefore shares a trait or feature with nothing. To distinguish groups of suffixes that are different with respect to triggering LVL is more theoretically problematic in a shortening approach as their most important trait is overtness.

In the case of lengthening analyses, an intuitive idea can be formalized for Hungarian Low Vowel Lengthening: it is the addition of an extra morpheme (i.e., the suffix)

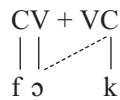
---

1 Cross-linguistically, this would not be unprecedented. McCarthy (2005: 11), for example, argues that this is in fact the only way to derive such alternations, given that “[there] may be other lengthening contexts, but presuffixal position does not seem to be one of them.”

that triggers the lengthening. This gives us the opportunity of explaining it – in, for example, an autosegmental framework – by the suffix providing extra timing units that the stem-final low vowel can spread to. Such an approach places the structural difference in the suffix, allowing for a more practical differentiation between suffixes that trigger LVL and those that do not. In the following, an approach like this, our approach, will be presented.

#### 4. EMPTY V IN THE SUFFIX

In this analysis suffixes triggering LVL begin with an empty vowel position. The empty V-position enables the melody of the word final short low vowel to spread and that is how it lengthens. The underlying form of the stem contains a single word-final /ɔ/ or /ε/ and when it is lengthened, its long form will be [a:] or [e:] (as proposed in Section 1). This is illustrated in the following figure with a simple case—the word *fák* [fa:k] ‘trees PL NOM’. This example gives a good insight into the structure of most words showing Low Vowel Lengthening.



##### 4.1 Lowering Stems

Lowering Stems in this framework can be represented as stems with a floating element just like in Rebrus (2000). It can be stipulated that *gáz* ‘gas’, which is a regular stem, and *ház* ‘house’, which is a Lowering Stem, both have a CV structure of the form CVVC, but *ház* also has a floating A archiphoneme. (This A is a low vowel not specified for backness, as Lowering Stems may take either ɔ or ε as a linking sound but that is in all cases deducible from the harmonic properties of the stem).



The figure above shows the same suffix (the Accusative Suffix) with three different types of stems. The first word is *kutyát* ‘dog, acc.’ in (i.) where Low Vowel Lengthening takes place as previously described. The second example in (ii.) is *gázt* ‘gas, acc.’ which is a completely regular noun and does not take any linking vowels. The third one in (iii.) is a Lowering Stem and the floating A of the stem attaches to the empty vowel position at the beginning of the suffix.

This is a case where the suffix chosen does not take a linking vowel with regular stems. In the example in the next figure, however, there is the Plural Suffix *-k* that always

appears as a vowel + consonant string if attached to a word that ends in a consonant. The structure of the suffix proposed by this analysis is the following: it is made up of a VC string to whose consonant position a 'k' is linked, but whose V-position is empty.

Furthermore, it has a floating vowel which is a rounded mid vowel underspecified for any other melodic property, marked by an O. The first example (*kutyák* 'dog PL NOM') shows a word in whose Plural form LVL can be observed. The other two cases (*gázok* 'gas PL NOM' and *házak* 'house PL NOM') are examples of regular and Lowering Stems, respectively. In *gázok* the floating O occupies the vowel position and is articulated (as [o]). In *házak*, the difference is the floating element of the stem and, as it becomes associated to the suffix's empty vowel position, the O will remain stranded and unpronounced.



The order that specifies which elements will be linked at the end of the derivation and therefore be pronounced on the surface seems complicated at first sight. Some possible rules and principles are formulated in (2) and (3).

- (2) Floating elements do not link to empty positions if that would create hiatus.
- (3) If the stem has a floating vowel and the suffix has an empty vowel position, the floating stem vowel will associate to it.

However, these separate rules and conditions in (2) and (3) prove to be an unnecessary complication. If this autosegmental analysis is enriched by specifying the direction of mapping as left-to-right, these conditions are automatically formulated in a much simpler and more uniform way.

## 4.2 Vowel Harmony

In this approach the distinction between suffixes triggering LVL and those not triggering it is made on a purely structural basis. The only suffixes capable of triggering LVL begin with an empty vowel slot, to which the morpheme final low vowel of the stem can spread. As there is no domain boundary stipulated – not even in cases of suffixes like *-sÁg* that do not trigger LVL – there is no reason why Vowel Harmony would be blocked in any way. Therefore, in this analysis, there is no interaction between Vowel Harmony and Low Vowel Lengthening.

## 4.3 Pros and Cons

Probably the main strength of this theory is that it explains the alternation by a structural difference in the suffix and not the stem. Looking at the distribution of LVL, it

seems that whether it takes place or not depends much more on the suffix than on the stem (once the stem provides a potentially sufficient environment – i.e., it ends in a low vowel). On the other hand, suffixes seem to trigger or not trigger it arbitrarily. Therefore it only seems logical to suppose two different structure-types for the suffixes and not for the stems. Should this analysis prove right even if it is extended to the whole vowel system, the theory itself is capable of providing a simple but powerful solution to Low Vowel Lengthening. The question of LVL intervening in Vowel Harmony does not even arise in this analysis.

However, it is clearly a weakness of the theory that it has been created only with respect to the low vowels of Hungarian. Testing how this structure-system of suffixes works with the other vowels and consonants of Hungarian is a topic left for future research.

## 5. CONCLUSION

The arguments cumulated in this paper corroborate a lengthening approach to Hungarian Low Vowel Lengthening. Moreover, the phenomenon can be effectively analyzed as a lengthening process, which simultaneously satisfies both requirements established in Section 2. First, since it makes a marked structure in the suffix the reason for LVL, it explains why certain suffixes do not trigger it – non-triggering suffixes do not contain the suffix-initial empty V position. Second, the empty-V approach can also integrate the class of Lowering Stems in Hungarian by supposing that they contain a floating vowel, a frequent component in the analyses of these stems. Such floating segments can then be associated to the empty V position at the beginning of certain suffixes. However, further compatibility of the analysis with other phonological phenomena in Hungarian is yet to be investigated.

## References

- ABONDOLO, Daniel M. (1988) *Hungarian Inflectional Morphology*. Budapest: Akadémiai Kiadó.
- KIPARSKY, Paul (1968) *How abstract is phonology?* Bloomington: Indiana University Linguistics Club.
- McCARTHY, John (2005) *The length of stem-final vowels in Colloquial Arabic* [manuscript]. Available at: [http://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1084&context=linguist\\_faculty\\_pubs](http://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1084&context=linguist_faculty_pubs)
- NÁDASDY, Ádám/Péter SIPTÁR (1994) “A magánhangzók.” In: F. Kiefer (ed.), *Strukturális magyar nyelvtan 2. Fonológia*. Budapest: Akadémiai Kiadó, 42–182.
- NOVÁK, Attila (1999) *Inflectional paradigms in Hungarian. The conditioning of suffix- and stem-alternations*. MA Thesis, Eötvös Loránd University, Budapest.
- REBRUS, Péter (2000) “Morfofonológiai jelenségek.” In: F. Kiefer (ed.), *Strukturális magyar nyelvtan III. Alaktan*. Budapest: Akadémiai Kiadó, 763–947.
- SIPTÁR, Péter/Miklós TÖRKENCZY (2000) *The Phonology of Hungarian (The Phonology of the World’s languages)*. New York: Oxford University Press Inc.

Summary  
LOW VOWEL “LENGTHENING” IN HUNGARIAN

This paper examined the topic of Low Vowel Lengthening in Hungarian, which is a term describing the short-long alternation that low vowels show. After an introduction of the vowel system and phonotactics of the language, two main criteria were identified that an analysis of LVL has to satisfy: (i) being able to explain suffixes that do not trigger LVL, yet interact with the stem and (ii) being a close-fitting model for other phenomena related to linking vowels, as well (the need for the latter rose from a functional motivation).

From the two possible groups of analyses (lengthening and shortening approaches), it was lengthening that proved to be a more suitable account for the phenomena. Examples of both were given with explanation and evaluation on the two criteria. Finally, the empty-V approach suggested in this paper was also tested against these criteria.

**Keywords:** Hungarian phonology, Low Vowel Lengthening, Lowering Stems, phonology

Povzetek  
»DALJŠANJE« NIZKIH SAMOGLASNIKOV V MADŽARŠČINI

Članek obravnava daljšanje nizkih samoglasnikov v madžarščini, tj. pojav premenjavanja kratke in dolge oblike, ki nastopi pri nizkih samoglasnikih. V uvodu je predstavljen samoglasniški sestav in fonotaktična pravila jezika. Sledi opis dveh glavnih kriterijev, ki jima mora zadostiti analiza daljšanja nizkih samoglasnikov: i) analiza mora pojasniti pripone, ki ne sprožijo daljšanja kljub interakciji z osnovo; ii) analiza mora biti model za druge pojave, ki so sorodni s povezovalnimi samoglasniki.

Izmed dveh možnih pristopov k analizi pojava (daljšanje ali krajšanje) je daljšanje tisto, ki bolj uspešno razloži pojav. V članku sta podana oba pristopa skupaj z razlago in njunim vrednotenjem glede na zgoraj omenjena kriterija. Članek na podoben način obravnava tudi t.i. pristop ničtega samoglasnika.

**Ključne besede:** madžarska fonologija, daljšanje nizkih samoglasnikov, fonologija, osnove z nižanjem







## INDEXICAL DEMONSTRATIVES AND IDENTIFICATIONAL FOCUS IN HUNGARIAN

### 1. INTRODUCTION

According to the traditional view, the choice of demonstratives (proximal *this* vs. distal *that*) is determined by the relative distance from the speaker. However, this view has been challenged by various authors (Piwek et al. 2008; Enfield 2009; Diessel 2012). For instance, Piwek et al. (2008) suggested that in Dutch acceptability should replace distance as a decisive factor influencing the choice of demonstratives. Recently, Tóth et al. (2014) examined the factors influencing the use of demonstratives in Hungarian, Dutch and English. The results show that the distribution of demonstratives (the number of proximal demonstratives and distal demonstratives) is significantly different in neutral and contrastive contexts when the entities being referred to are close to the speaker. This paper collects further data about the use of Hungarian demonstratives in contrastive contexts via investigating a special subtype of contrastive contexts marked by identificational focus. If the distribution of demonstratives in identificational focus is the same as their distribution in other types of contrastive contexts, then experimental evidence will support the theoretical claim about the contrastive nature of identificational focus.

### 2. THE USE OF HUNGARIAN DEMONSTRATIVES

In Hungarian, there are two types of demonstratives, *ez/ezek* ‘this/these’ are proximal, whereas *az/azok* ‘that/those’ are distal demonstratives. Regarding their uses, indexical demonstratives are those that are accompanied by a pointing gesture. Levinson (2004) divides indexical uses into two subcategories: non-contrastive (1) and contrastive uses (2-3), as illustrated below.

---

\* toth.eniko@arts.unideb.hu

\*\* csatar.peter@arts.unideb.hu

- (1) *Ez a dinnye nagyon finom.*  
 this the melon very tasty  
 DEM.PROX.NOM.SG NOM  
 ‘This melon is delicious.’
- (2) *Az a futó nyert, és nem ez.*  
 that the runner win and not this  
 DEM.DIST.NOM.SG NOM 3SG.PAST DEM.PROX.NOM.SG  
 ‘That runner has won the race and not this one.’
- (3) *Ezt a dinnyét kérem.*  
 this the melon want  
 DEM.PROX.ACC.SG ACC 1SG.PRESENT  
 ‘I want this melon.’

Extending Levinson’s view and relying on the analyses of Chafe (1994), É. Kiss (1998, 2002) and Kaiser (2011), contrastive contexts are defined as follows:

- (i) physical context: no conditions;
- (ii) epistemic context: the entities are activated in the discourse and they are highly accessible for the participants;
- (iii) linguistic context: contrastiveness is explicitly indicated linguistically, for instance by using a coordinating conjunction with a contrastive sense, e.g., *but*, or by prosodic prominence;
- (iv) social context: not relevant.

Contexts that do not satisfy the definition above are labelled as neutral.

## 2.1 Hungarian Demonstratives and Identificational Focus

Identificational focus in Hungarian is marked by stress and the focussed constituent moves into a preverbal position. Moreover, if the verb contains a preverb (see the examples below), it will leave its verb and move into a position that is immediately after the verb. The example in (4) illustrates a neutral sentence (i.e., neutral context in the experiment to be presented), whereas in (5) the NP containing the indexical demonstrative is in identificational focus (i.e., contrastive contexts later on).

- (4) *Meg-veszem azt a könyvet.*  
 buy that the book  
 PREVERB-1SG.PRES DEM.DIST.ACC.SG ACC  
 ‘I’ll buy that book.’
- (5) *Azt a könyvet veszem meg.*  
 that the book buy  
 DEM.DIST.ACC.SG ACC 1SG.PRES PREVERB  
 ‘It’s that book that I’ll buy.’

There are several theories regarding the syntactic and semantic characteristics of identificational focus. Syntactic theories are concerned with explaining how the movement is triggered, while semantic approaches concentrate on the nature of the exhaustive interpretation of identificational focus. Here only the latter approaches will be described briefly.

The function of Hungarian identificational focus is defined by É. Kiss (1998, 2002) as follows: “The focus represents a proper subset of the set of contextually or situationally given referents for which the predicate phrase can potentially hold; it is identified as the exhaustive subset of this set for which the predicate phrase holds.” (É. Kiss 2002: 78)

In general, all theories accept that there is an exhaustive interpretation associated with Hungarian identificational focus. Proponents of the standard analysis (É. Kiss 1998, 2002; Szabolcsi 1981) claim that exhaustivity is a semantic feature, i.e., exhaustivity is part of the truth-conditions of sentences with identificational focus. This view has been challenged recently, for instance by Wedgwood (2005), who claims that exhaustivity is an implicature in a relevance theoretical framework, and by Balogh (2009), who argues that exhaustivity is an obligatory implicature in an inquisitive semantic framework. Empirical studies have also questioned the standard semantic feature theory (for details see Kas and Lukács 2013; Geröcs et al. 2014). In her more recent papers É. Kiss (2004, 2006) analyses identificational focus as a specificational predicate and argues that exhaustivity is implied. As it is clear from this brief theoretical overview, there are controversial results regarding the treatment of identificational focus and the status of exhaustivity. For the purposes of the present paper we adopt Balogh’s (2009) view, who argues that the exhaustive interpretation is obligatory, provided that the verb contains a preverb and it moves into a postverbal position (Balogh 2009: 139).

Furthermore, É. Kiss (1998) also argues that Hungarian identificational focus can be [ $\pm$ contrastive]. More specifically, it is [+contrastive] “if it operates on a closed set of entities whose members are known to the participants of the discourse. [...] In this case, the identification of a subset of the given set also identifies the contrasting complementary subset” (É. Kiss 1998: 267). The study reported here is aimed at investigating the behaviour of indexical demonstratives in identificational focus and to see whether empirical data can also support this view.

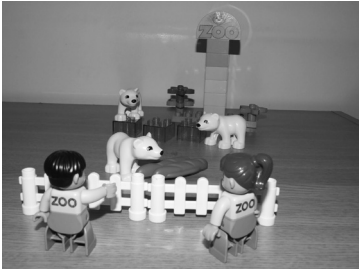
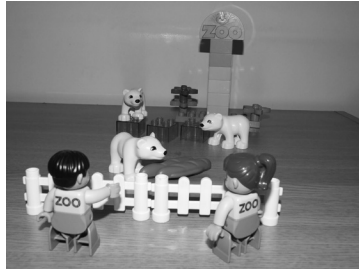
### 3. THE EXPERIMENT

The aim of our experiment is twofold. On the one hand, it wants to reinforce the role of distance in neutral contexts with a different method. On the other hand, the second hypothesis tests whether utterances with distal demonstratives in identificational focus receive higher ratings than utterances with indexical demonstratives in neutral non-focus position. If the findings support the second hypothesis, then empirical evidence will be provided in favour of the claim that identificational focus is a syntactically marked subtype of contrastive contexts.

### 3.1 Materials and Methods



To explore the hypotheses above we constructed an online questionnaire, which included both neutral and contrastive contexts. In order to ensure maximum reliability of our test we created rich contexts (Meibauer 2012). The task of the participants was to evaluate the acceptability of a reply to a question with respect to a context represented by a picture on a four-point Likert scale (forced choice method). In the picture Lego DUPLO figures displayed a zoo scenario. For the sake of the distance hypothesis, the participants of the mini-dialogue (including the speaker) also appeared in the picture, and a pointing gesture accompanying the use of a certain demonstrative was emphasized in two different ways: (i) it was explicitly mentioned in the dialogue; (ii) the Lego figure depicting the speaker also used a pointing gesture. The location of the referred entities was also clear from the context (all entities being referred to were close to the speaker).

Figure 1: A –contrastive, +proximal test item

<p>Az új gondozót körbevezetik az állatkertben. A jegesmedvéknél az új gondozó megkérdezi: – Mit csináltál az előbb a medvéknél?</p>  <p>– Ezt a jegesmedvét (rámutat) odacsaltam a száraz kenyérhez.</p> <ul style="list-style-type: none"> <li><input type="radio"/> 1 teljesen elfogadhatatlan</li> <li><input type="radio"/> 2 kevésbé elfogadhatatlan</li> <li><input type="radio"/> 3 kevésbé elfogadható</li> <li><input type="radio"/> 4 teljesen elfogadható</li> </ul>	<p>“The new zoo tender is being shown round in the zoo. He/she asks at the polar bear zone.” “What have you been doing at the polar bears?”</p>  <p>“I have lured this bear (pointing to the bear) to the dry baguette.”</p> <ul style="list-style-type: none"> <li><input type="radio"/> 1 totally unacceptable</li> <li><input type="radio"/> 2 slightly unacceptable</li> <li><input type="radio"/> 3 slightly acceptable</li> <li><input type="radio"/> 4 totally acceptable</li> </ul>
--	--

Moreover, in contrastive contexts the exhaustive interpretation was also prompted with the help of a supportive stimulus which highlighted the difference between the elements identified by the focus and those of the complementary set. In Figure 2 the bandage on the giraffe’s leg helps to pick out the entity being referred to.

Figure 2: A +contrastive, –proximal test item

<p>A két gondozó délután az állatorvoshoz készül a zsiráfokkal. - Melyik zsiráfot visszük el délután az állatorvoshoz?</p>  <p>- Azt a zsiráfot (rámutat) visszük el.</p> <ul style="list-style-type: none"> <li>○ 1 teljesen elfogadhatatlan</li> <li>○ 2 kevésbé elfogadhatatlan</li> <li>○ 3 kevésbé elfogadható</li> <li>○ 4 teljesen elfogadható</li> </ul>	<p>The two zoo tenders are going to take the giraffes to the vet in the afternoon. “Which giraffe are we taking to the vet in the afternoon?”</p>  <p>“It’s that giraffe (pointing to the giraffe) that we are taking.”</p> <ul style="list-style-type: none"> <li>○ 1 totally unacceptable</li> <li>○ 2 slightly unacceptable</li> <li>○ 3 slightly acceptable</li> <li>○ 4 totally acceptable</li> </ul>
---	--

We used a 2x2 within–subjects design, with the factors ( $\pm$ contrastive and  $\pm$ proximal). For instance, the test item in Figure 2 represents a (+contrastive, –proximal) item. The test included 4 items in each condition and 8 filler contexts. The items were presented in a pre-set random order which was the same for each participant. In each of the test items preverb-verb constructions were used, and identificational focus was marked explicitly by moving the preverb into a postverbal position,

36 participants, all native speakers and students at the University of Debrecen, took part in the experiment. Their average age was 22 years. The participants did not know about the purpose of the experiment.

### 3.2 Results and Discussion

To test the first hypothesis (the role of distance in neutral contexts) we compared (–contrastive,  $\pm$ proximal) conditions using the sign test. As expected, there was a significant difference between the ratings of utterances (sign test,  $z = -2.00$ ,  $p < 0.05$ ), i.e., utterances with proximal demonstratives were preferred when referring to entities that were close to the speaker (see Table 1 and Table 2, Figure 3).

Table 1: Distribution of ratings in the –contrastive, +proximal condition

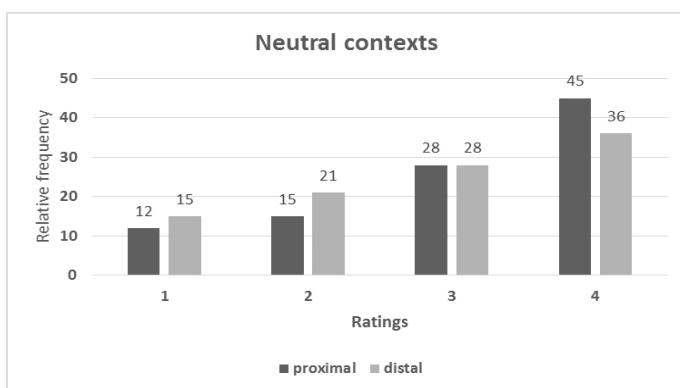
Ratings	Frequency	Relative frequency (%)
1	17	12 %
2	22	15 %
3	41	28 %
4	64	45 %
<b>Total</b>	144	100 %

Table 2: Distribution of ratings in the –contrastive, –proximal condition

Ratings	Frequency	Relative frequency (%)
1	21	15 %
2	30	21 %
3	41	28 %
4	52	36 %
<b>Total</b>	144	100 %

Hence, as Figure 3 shows, the results reinforced the role of distance: native speakers preferred indexical proximal demonstratives referring to entities that were close to the speaker (cf. Tóth et al. 2014). Coventry et al. (2008) and Luz and Van der Sluis (2011) drew similar conclusions for English, and English, Dutch and Portuguese, respectively. From a methodological point of view, the present results provide converging evidence for the crucial role of distance in neutral contexts.

Figure 3: Relative frequency of ratings in the –contrastive, ±proximal conditions



To test the second hypothesis (neutral vs. contrastive contexts) we used the Friedman test, which yielded a significant result across the four conditions ( $\chi^2(3) = 21.85$ ,  $p < 0.01$ ). Post-hoc tests (Wilcoxon signed rank tests) detected a significant difference (using a Bonferroni-adjusted significance level) between the acceptability ratings of (–proximal) sentences in contrastive and neutral contexts ( $z = -4.25$ ,  $p < 0.01$ ), i.e., in contrastive contexts, which have been marked by identificational focus, distal demonstratives were judged to be more acceptable than in neutral contexts (see Table 3 and Table 4, Figure 4). It also has to be noted here that from a descriptive statistical perspective this condition was the only one when the modus, i.e., the most frequently selected value, was the highest one, 4. This means that participants found the utterances with distal demonstratives in identificational focus the most acceptable.

Table 3: Distribution of ratings in the –contrastive, –proximal condition

Ratings	Frequency	Relative frequency (%)
1	21	15 %
2	30	21 %
3	41	28 %
4	52	36 %
<b>Total</b>	144	100 %

Table 4: Distribution of ratings in the +contrastive, –proximal condition

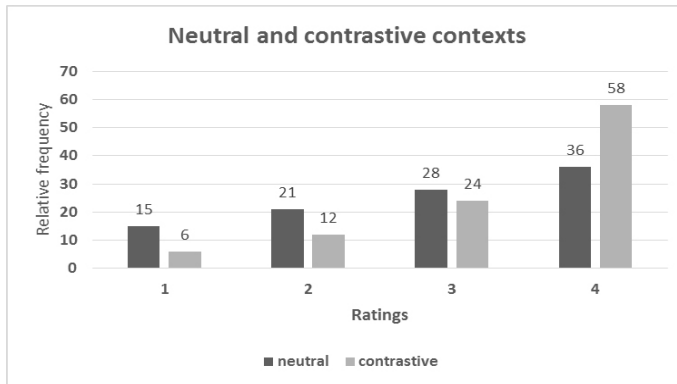
Ratings	Frequency	Relative frequency (%)
1	9	6 %
2	18	12 %
3	34	24 %
4	83	58 %
<b>Total</b>	144	100 %

There is no significant difference between the ratings of (+proximal) utterances in neutral and contrastive contexts. Since the entities being referred to were always close to the speaker, the results are in line with the traditional approach; we expected high ratings in both types of context.

Similarly, there is no significant difference between the ratings of (+contrastive) utterances containing proximal and distal demonstratives in identificational focus. This suggests that in contrastive contexts not only distal demonstratives, but also proximal demonstratives are acceptable, supposing that the entity being referred to is *close* to the speaker. If only distal demonstratives could be used in contrastive contexts, a significant difference should have been found between (+contrastive) and ( $\pm$ proximal)



Figure 4: Relative frequency of ratings in the  $\pm$ contrastive,  $-$ proximal conditions



conditions. At the same time it has been proved again that distance on its own cannot explain the use of indexical demonstratives in contrastive contexts. The results above are in accordance with Levinson’s (2000) view, who argues that demonstratives in English form a Q-contrast  $\langle$ this, that $\rangle$ , which means that the use of *this* has to satisfy the criterion of proximity, while the use of *that* is not restricted in this sense. “This predicts that *that* has a wide distribution, potentially overlapping with *this*, as indeed seems to be the case” (Levinson 2000: 94). Our results support the same view regarding the use of Hungarian indexical demonstratives.

#### 4. CONCLUSION

To conclude, first, the results provide converging evidence and reinforce that distance plays a crucial role in neutral contexts (when the entities being referred to are close to the speaker). Second, utterances with distal demonstratives in identificational focus received significantly higher ratings. Tóth et al. (2014) also showed that distal demonstratives in contrastive contexts are preferred to proximal demonstratives. Therefore, the theoretical claim that identificational focus forms a subtype of contrastive contexts has been reinforced by the results presented above. More specifically, our findings provide empirical evidence in favour of É. Kiss’s (1998) theory about the contrastive nature of identificational focus, at least under the condition that identificational focus is explicitly marked by moving the preverb to a postverbal position.

#### Acknowledgements

This research was supported by the National Research, Development and Innovation Office (NKFIH), grant No. K 111918.

## References

- BALOGH, Kata (2009) *Theme with variations: a context-based analysis of focus*. Amsterdam: ILLC Dissertation Series.
- CHAFE, Wallace (1994) *Discourse, consciousness, and time: the flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.
- COVENTRY, Kenny R./Berenice VALDÉS/Alejandro CASTILLO/Pedro GUIJARRO-FUENTES (2008) "Language within your reach: near-far perceptual space and spatial demonstratives." *Cognition* 108, 889–895.
- DIESSEL, Holger (2012) *Deixis and demonstratives*. In: C. Maienborn/K. von Heusinger/P. Portner (eds), 2407–2431.
- EMONDS, Joseph/Markéta JANEBOVÁ (eds) (2014) *Language use and linguistic structure. Proceedings of the Olomouc Linguistics Colloquium 2013*. Olomouc: Palacky University.
- ENFIELD, Nick J. (2009) *The anatomy of meaning. Speech, gesture, and composite utterances*. Cambridge: Cambridge University Press.
- É. KISS, Katalin (1998) "Informational focus vs. identification focus." *Language* 74/2, 245–273.
- É. KISS, Katalin (2002) *The syntax of Hungarian*. Cambridge: Cambridge University Press.
- É. KISS, Katalin (2004) *Identificational focus and informational focus revisited*. In: L. Hunyadi/Gy. Rákosi/E. Tóth (eds), 59–72.
- É. KISS, Katalin (2006) *Focussing as predication*. In: V. Molnár/S. Winkler (eds), 169–196.
- FINKBEINER, Rita/Jörg MEIBAUER/Petra B. SCHUMACHER (eds) (2012) *What is a context? Linguistic approaches and challenges*. Amsterdam: John Benjamins.
- GARDENT, Claire/Kristina STRIEGNITZ (eds) (2011) *Proceedings of the 13th European workshop on natural language generation*. Stroudsburg (PA): Association for Computational Linguistics.
- GERÓCS, Mátyás/Anna BABARCZY/Balázs SURÁNYI (2014) *Exhaustivity in focus: experimental evidence from Hungarian*. In: J. Emonds/M. Janebová (eds), 181–194.
- HORN, Laurence R./Gregory WARD (eds) (2004) *The handbook of pragmatics*. Oxford: Blackwell.
- HUNYADI, László/György RÁKOSI/Enikő TÓTH (eds) (2004) *Proceedings of the 8th symposium on logic and language*. Debrecen: University of Debrecen.
- KAISER, Elsi (2011) "Salience and contrast effects in reference resolution: The interpretation of Dutch pronouns and demonstratives." *Language and Cognitive Processes* 26/10, 1587–1624.
- KAS, Bence/Ágnes LUKÁCS (2013) "Focus sensitivity in Hungarian adults and children." *Acta Linguistica Hungarica* 60/2, 217–245.
- LEVINSON, Stephen C. (2000) *Presumptive meanings: the theory of generalized conversational implicature*. Cambridge (MA): MIT Press.

- LEVINSON, Stephen C. (2004) *Deixis and pragmatics*. In: L. R. Horn/G. Ward (eds), 97–121.
- LUZ, Saturnino/Ielka VAN DER SLUIS 2011. *Production of demonstratives in Dutch, English and Portuguese dialogues*. In: C. Gardent/K. Striegnitz (eds), 181–186.
- MAIENBORN, Claudia/Klaus VON HEUSINGER/Paul PORTNER (eds) (2012) *Semantics. An international handbook of natural language meaning*. Berlin: Mouton de Gruyter.
- MEIBAUER, Jörg (2012) *What is a context? Theoretical and empirical evidence*. In: R. Finkbeiner/J. Meibauer and P. B. Schumacher (eds), 9–32.
- MOLNÁR, Valéria/Susanne WINKLER (eds) (2006) *The architecture of focus*. Berlin: Mouton de Gruyter.
- PIWEK, Paul/Robbert-Jan BEUN/Anita CREMERS. (2008) “‘Proximal’ and ‘distal’ in language and cognition: evidence from deictic demonstratives in Dutch.” *Journal of Pragmatics* 40, 694–718.
- SZABOLCSI, Anna (1981) “Compositionality in focus.” *Folia Linguistica* 15, 141–161.
- TÓTH, Enikő/Péter CSATÁR/Arina BANGA (2014) *Exploring Hungarian and Dutch gestural demonstratives*. In: L. Veselovská/M. Janebová (eds), 607–625.
- VESELOVSKÁ Ludmila/Markéta JANEBOVÁ (eds) (2014) *Complex visibles out there. Proceedings of the Olomouc Linguistics Colloquium 2014: language use and linguistic structure*. Olomouc: Palacky University.
- WEDGWOOD, Daniel (2005) *Shifting the focus. From static structures to dynamics of interpretation*. Amsterdam: Elsevier.

### Summary

#### INDEXICAL DEMONSTRATIVES AND IDENTIFICATIONAL FOCUS IN HUNGARIAN

This paper presents the results of an experiment regarding the use of Hungarian indexical demonstratives where it is shown that the use of indexicals depends on the nature of the context. More specifically, the use of indexical demonstratives is explored in neutral contexts and in a subtype of contrastive contexts, where contrastiveness is indicated by Hungarian identificational focus. On the one hand, the results reinforce our previous findings that distance is a crucial factor in neutral contexts. On the other hand, it is revealed that utterances with distal demonstratives in identificational focus receive higher ratings on a Likert-scale than utterances with distal demonstratives in a neutral (non-focus) position. These results provide evidence in favour of the contrastive nature of identificational focus in Hungarian, at least under the condition that identificational focus is explicitly marked by moving the preverb to a postverbal position.

**Keywords:** experimental pragmatics, identificational focus, indexical demonstratives

Povzetek  
INDEKSIALNI KAZALNI ZAIMKI IN IDENTIFIKACIJSKI  
FOKUS V MADŽARŠČINI

V članku predstavimo rezultate eksperimenta o indeksialnih kazalnih zaimkih v madžarščini, ki pokaže, da je raba indeksialnih zaimkov odvisna od narave konteksta. Raba indeksialnih kazalnih zaimkov smo preverili v nevtralnih kontekstih in v podtipu kontrastivnih kontekstov, kjer je kontrastivnost izražena s pomočjo madžarskega identifikacijskega fokusa. Rezultati na eni strani potrjujejo naša predhodna dognanja, da je oddaljenost ključni faktor v nevtralnih kontekstih. Na drugi strani je iz rezultatov razvidno, da so izreke z distalnimi kazalnimi zaimki v identifikacijskem fokusu na Likertovi lestvici uvrščene višje kakor izreke z distalnimi kazalnimi zaimki v nevtralnem (nefokusnem) položaju. Rezultati kažejo na kontrastivno naravo identifikacijskega fokusa v madžarščini – vsaj pod pogojem, da je identifikacijski fokus eksplicitno označen s premikom iz predlagolskega v zaglagolski položaj.

**Ključne besede:** eksperimentalna pragmatika, identifikacijski fokus, indeksialni kazalni zaimki





## AMBIGUOUS ADJECTIVES IN FRENCH: THE CASE OF “GROS” WHEN COMBINED WITH DEVERBAL NOUNS

### 1. INTRODUCTION

The problem of adjectival ambiguity is well-known, and many formal semanticists have tried to account for it since the early days of the discipline (Siegel 1976; Larson 1998). The fact that adjectives can have different readings depending on their position within the DP (for example, prenominal or postnominal) has also been analysed (see Cinque 2010, among others).

In this paper, I will focus on the adjective *gros* in French (“big”, “fat”). This adjective is not interpreted in the same way when it appears before or after an *-eur* deverbale noun (see the examples in (1)).

- |                              |                          |
|------------------------------|--------------------------|
| (1) a) <i>Un gros fumeur</i> | b) <i>Un fumeur gros</i> |
| a GROS smoker                | a smoker GROS            |
| ✓ ‘A heavy smoker’           | * ‘A heavy smoker’       |
| ? ‘A fat smoker’             | ✓ ‘A fat smoker’         |

My goal is to explain the French facts by giving a compositional analysis of the phenomenon presented in (1). In the literature, two sources have been diagnosed in the literature for the ambiguity of DPs: the first is the adjective (“blame *Adj*”, see Siegel 1976; Despić and Sharvit 2008), the second the noun (“blame *NP*”, see Larson 1998). I will propose that another option, namely, “blame both”, accounts for the French data better.

The paper is outlined as follows. In section 1, I present the standard classification of adjectives into at least intersective modifiers and non-intersective modifiers. Then I discuss data from French, demonstrating the ambiguity of the adjective “*gros*”, and give generalisations that can be drawn from the data. In section 2, I review three previous analyses of ambiguous adjectives (Siegel 1976; Larson 1998; Winter and Zwarts 2012a,b). In section 3, I propose an analysis for DPs containing a deverbale noun and the adjective “*gros*” (both in the pre- and the postnominal position). The conclusion is concerned with open issues and further research.

---

\* lucastual@unige.ch

## 2. AMBIGUOUS ADJECTIVES

### 2.1 Classification of Adjectives

One possible classification of adjectives (see Partee 1995, a.o.) divides them into two categories: intersective (such as “*red*”), and non-intersective (such as “*skillful*”). The non-intersective class is often also subdivided into other categories (plain non-subjective, privative non-subjective, see Kamp and Partee 1995), but this fine-grained classification has been refuted, notably by Partee (2010). For the purposes of this paper, the distinction between intersective and non-intersective modifiers is sufficient, and thus I will not consider subdivisions of the non-intersective category.

When an adjective is intersective, as is the case with *red*, the denotation of an NP containing a noun and this kind of adjective is the *intersection* of the semantic value of the noun and the semantic value of the adjective, as described formally in (2).

$$(2) \quad [[Adj_{intersective} Noun]] = [[Adj_{intersective}]] \cap [[Noun]]$$

In a case such as (2), it is possible to say that everything that is an “ $Adj_{intersective} Noun$ ” is both an “ $Adj_{intersective}$ ” and a “ $Noun$ ”. Let me illustrate this point with the following inference:

- (3) *This is a red table.*  
a)  $\rightarrow$  *This is a table.*  
b)  $\rightarrow$  *This is red.*

Non-intersective adjectives do not follow the pattern just presented. A classic example of a non-intersective adjective is “*skillful*”. If we reproduce the inference in (3) with the adjective “*skillful*”, the entailment pattern is not similar:

- (4) *This man is a skillful surgeon.*  
a)  $\rightarrow$  *This man is a surgeon.*  
b)  $\nrightarrow$  *This man is skillful.*

Informally speaking, it is possible to say that someone is skillful *as a surgeon*, but that does not entail that this person is skillful in *general*. Hence, the right way to describe a non-intersective adjective is to say that the denotation of a NP containing a noun and a non-intersective adjective is *not* the intersection of the semantic value of the noun and the semantic value of the adjective:

$$(5) \quad [[Adj_{non-intersective} Noun]] \neq [[Adj_{non-intersective}]] \cap [[Noun]]$$

These kind of adjectives can be seen as having a more complex semantic type than intersective adjectives.

Having discussed the standard classification of adjectives, I will now present novel French data in which a specific adjective is ambiguous between an intersective and a non-intersective reading.

## 2.2 The French Data

The French data presented in this section will be analysed in section 3. I will describe the interaction between the adjective “*gros*” and an *-eur* deverbal noun. (6–9) demonstrates that the interpretation of a DP containing both of these elements depends on whether or not the adjective is in the prenominal or the postnominal position.

- |  |  |
|--|--|
| (6) a) <i>Un gros fumeur</i><br>a GROS smoker<br>✓ ‘A heavy smoker’<br>? ‘A fat smoker’              | b) <i>Un fumeur gros</i><br>a smoker GROS<br>* ‘A heavy smoker’<br>✓ ‘A fat smoker’              |
| (7) a) <i>Un gros dormeur</i><br>a GROS sleeper<br>✓ ‘Someone who sleeps a lot’<br>? ‘A fat sleeper’ | b) <i>Un dormeur gros</i><br>a sleeper GROS<br>* ‘Someone who sleeps a lot’<br>✓ ‘A fat sleeper’ |
| (8) a) <i>Un gros buveur</i><br>a GROS drinker<br>✓ ‘A heavy drinker’<br>? ‘A fat drinker’           | b) <i>Un buveur gros</i><br>a drinker GROS<br>‘A heavy drinker’<br>✓ ‘A fat drinker’             |
| (9) a) <i>Un gros mangeur</i><br>a GROS eater<br>✓ ‘A big eater’<br>? ‘A fat eater’                  | b) <i>Un mangeur gros</i><br>a eater GROS<br>* ‘A big eater’<br>✓ ‘A fat eater’                  |

In all of the *a*) examples, the DP has a strongly preferred non-intersective reading: “*un gros fumeur*” will refer to a person who smokes a lot. Under this reading, the adjective is truly non-intersective, as is demonstrated by the failed entailment pattern in (10).

- (10) *Cet homme est gros fumeur.*  
This man is GROS smoker  
a. → *Cet homme est fumeur.*  
this man is smoker  
b. → *Cet homme est gros.*  
this man is fat

It is nonetheless possible to give an intersective interpretation to this kind of sentence, where “*un gros fumeur*” would refer to a person who smokes and who is fat.<sup>1</sup> However, when this very same adjective is in a postnominal position (see the *b*) examples), the DP *only* bears an intersective reading, where “*gros*” means “fat”.

Until now, I have only considered the interpretation of the adjective *gros* when it is used attributively. Let us consider now the use of this adjective in predicative constructions. (11) demonstrates that it only has an intersective reading.

1 The “?” next to the second free translation in *a*) examples states that the preferred reading when the adjective is prenominal is the non-intersective one.



- (11) *Ce fumeur est gros.*  
 this smoker is fat

Concerning another adjective that behaves in a similar way to “*gros*”, we consider the case of “*petit*” (“small”). It is important to note that the adjective “*gros*” can be replaced by its antonym, “*petit*”, to obtain the opposite reading. Hence, “*un petit fumeur*” will denote a person who smokes only a little, “*un petit dormeur*” a person who sleeps only a little, and so on. “*Petit*” is also interpreted intersectively in a postnominal position: “*un petit fumeur*” describes a smoker of short height.

It should be noted that there is a correlation between the adjective “*gros*” and the adverb “*beaucoup*” in all of the non-intersective examples. As a matter of fact, the sentences in (12) and (13) convey the same meaning.

- |   |   |
|---|---|
| <p>(12) <i>C'est un gros fumeur.</i><br/>         it-is a GROS smoker<br/>         ‘He’s a heavy smoker.’</p> | <p>(13) <i>Il fume beaucoup.</i><br/>         he smokes lot<br/>         ‘He smokes a lot.’</p> |
|---|---|

The non-intersective interpretation of “*gros*” seems to be quite systematic. A corpus research from two French corpora (CFPP200 and ESLO1/2, see Branca-Rosoff et al. 2012; and Eshkol-Taravella et al. 2011, respectively) has demonstrated several other uses of the adjective “*gros*” with *-eur* deverbal nouns, where the interpretation is always non-intersective for a prenominal adjective: “*gros garsilleur*” (“big wasteful person”), “*gros utilisateur de X*” (“big user of X”), “*gros lecteur*” (“big reader”), “*gros acheteur*” (“big buyer”), “*gros fournisseur*” (“big supplier”), and “*gros transporteur*” (“big transporter”).

### 2.3 Generalisations and Goals

On the basis of the data discussed in the previous section, I derive the following generalisations for French concerning adjectives in the same category as “*gros*”:

1. Intersective adjectives can surface in different positions when used attributively (prenominal or postnominal).
2. Non-intersective adjectives always occupy a specific position in the structure.
3. In a predicative position, the adjective always receives an intersective interpretation.

Now that I have introduced these generalisations, the goals that I aim to achieve in this paper are listed below.

1. To observe the interaction of French *-eur* nominals with ambiguous adjectives.
2. To give a compositional analysis of the phenomenon.
3. In the literature, the discussion about ambiguous adjectives focused on deciding which part of the DP was responsible for the ambiguity: “Blame AP” (Siegel 1976; Despić and Sharvit 2008) or “Blame NP” (Larson 1998). I will argue in favor of a third possibility which is necessary to capture the ambiguity: “Blame both”.

Before presenting my analysis of ambiguous adjectives in French, I will discuss three analyses that have been proposed in the literature to account for similar data.

### 3. PREVIOUS ANALYSES OF AMBIGUOUS ADJECTIVES

In this section, I will first discuss the conflicting views of Siegel (1976) and Larson (1998) about adjectival ambiguity. The former advocates for an *A-analysis*, where the ambiguity results from the adjective only, while the latter proposes that an analysis in which the noun is responsible for the ambiguity is more plausible (*N-analysis*).

#### 3.1 Larson vs. Siegel

In the literature, many researchers have provided accounts of the ambiguity present in the sentence in (14).

(14) *Olga is a beautiful dancer.*

This sentence can be interpreted in an intersective manner, where what the speaker wants to convey is that Olga is both a dancer and a beautiful person. However, it also bears a non-intersective reading, where the meaning of the sentence is that Olga is beautiful *as a dancer*. In her dissertation, Siegel proposed a theory that could explain the ambiguity of (14), which has both an intersective and a non-intersective reading.<sup>2</sup> Siegel based her analysis on the fact that in Russian, the morphology of adjectives is such that they can appear in two different forms: one that is intersective only (the short-form) and another that is non-intersective (the long-form). This observation leads her to suggest that the adjective category – which has been considered to bring together every adjective – is to be conceived of as two separated classes of items that are syntactically and semantically distinct.

1. The first class contains adjectives that only appear in the predicative position, and are extensional and intersective, such as *aged*.
2. The second class contains adjectives that appear in the attributive or predicative position, and are intensional and non-intersective, such as *former*.

Siegel believes that this classification sheds light on the problem of adjectival ambiguity: while some adjectives appear only in one of these two classes, but not in the other (like *aged* or *former*), others like *beautiful* exist in two forms that are homophonous: one being in the first class, while the other is in the second class. In her dissertation, Siegel calls adjectives of this kind *doublets*.

From this assumption, the *doublet theory* for adjectives suggests that when two forms exist, one has an intersective semantic value, and the other a non-intersective

---

2 Unfortunately, I did not have access to Siegel's original document while writing this paper, thus the way her proposal is presented here is taken from indirect sources such as Larson (1998) and Morzycki (2013).

semantic value. This explains the data from English (for example, the sentence in (14)), in which ambiguity is due to the existence of two lexical entries for homophonous adjectives.

Contrary to Siegel, Larson (1998) develops a theory in which the ambiguity of the adjective *beautiful* in (14) is not due to the complexity of the adjective, but results more from the lexical properties of the noun it modifies (the so called *N-analysis*). In his analysis, Larson assumes that deverbal nouns like *dancer* involve event structure and a *genericity* operator. This analysis is based on the observation that a noun like *dancer* can be considered as derived from the verb *to dance*. According to Larson, at least in this structure, the verb *to dance* denotes the set of events of dancing.<sup>3</sup> Since the noun *dancer* possesses some event-structure, it can refer to an individual who habitually participates in an event of dancing. To achieve this, Larson postulates the presence of a genericity operator. The main point of this proposal is that, in a first step of the derivation, the noun *dancer* denotes sets of events, and, in a second step, it is combined with the GEN-operator, denoting sets of individuals at the end.

With this assumption, it is straightforward to account for the ambiguity of the adjective *beautiful* in the sentence *Olga is a beautiful dancer*: the adjective can either modify the event argument of the noun, giving rise to the non-intersective reading (“Olga dances beautifully”), or modify the individual argument of the noun (once the GEN-operator has applied) to give rise to the intersective reading (“Olga is a dancer and Olga is beautiful”).

Larson claims that this theory is an *N-analysis*, since he demonstrates that it is the complex structure of the noun that is responsible for the ambiguity of the sentence, and thus this ambiguity is supposed to be *structural* (as opposed to Siegel’s analysis that advocated for a *lexical* ambiguity).

### 3.2 Winter and Zwarts

In this last part of section 2, I present the theory proposed by (Winter and Zwarts 2012a,b) (henceforth W&Z), in which Larson’s proposal is further developed. Larson did not present a clear compositional theory in his paper, a gap that W&Z propose to fill.

W&Z introduce an operator that allows for the saturation of the event variable of a deverbal noun like *dancer*. During the derivation, the verb *dance* is first generated with type  $\langle v, et \rangle$ . The role of the *-er* morpheme is only morphological: it turns this verb into a noun syntactically. This morpheme is semantically vacuous: its semantic value can be seen as an identity function that takes a two-place predicate of type  $\langle v, et \rangle$  and results in the same kind of predicate (see (15)).

$$(15) \quad [[-er]] = \lambda P_{\langle v, et \rangle}. P$$

At this point, semantically *dancer* looks like a verb, since the event argument is still present (at this step of the derivation, the authors represent this category as *NPer*,

---

3 I will associate the type  $v$  to event arguments.

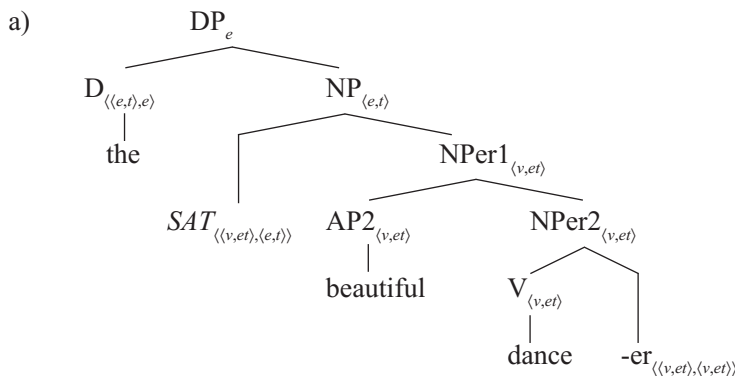
meaning that this constituent behaves syntactically like a noun, but that its event argument is not yet saturated).

W&Z assume, in a manner similar to Larson, that there is a covert operator (SAT) that saturates the event argument of the *NPer*, transforming it into a predicate of individuals (of type  $\langle e, t \rangle$ ). This operator has a semantic value similar to that of the GEN-operator used by Larson.

$$(16) \quad [[SAT]] = \lambda P_{\langle v, et \rangle} \cdot \lambda x_e \cdot \exists E [P(x)(E) \wedge [x \text{ habitually participates in } P\text{-events}]]$$

(17) demonstrates the derivation of the DP *the beautiful dancer* – on its non-interactive reading – and its semantic composition.

(17) *The beautiful dancer*



b)

$$\begin{aligned}
 [[-er]] &= \lambda P_{\langle v, et \rangle} \cdot P \\
 [[dance]] &= \lambda_{EV} \cdot \lambda x_e [dance(E) \wedge ag(E) = x] \\
 [[NPer1]] &= \lambda_{EV} \cdot \lambda x_e [dance(E) \wedge ag(E) = x \wedge beautiful(E)] \\
 [[SAT]] &= \lambda P_{\langle v, et \rangle} \cdot \lambda x_e \cdot \exists E [P(x)(E) \wedge [x \text{ habitually participates in } P\text{-events}]] \\
 [[NP]] &= \lambda x_e [dance(E) \wedge ag(E) = x \wedge beautiful(E) \wedge [x \text{ habitually participates in dancing-events}]]
 \end{aligned}$$

Even though W&Z’s proposal improved what Larson proposed by giving a more detailed compositional analysis of ambiguous DPs, I believe that the generic operator in their structure is not precise enough. As a matter of fact, according to both Larson and W&Z, the denotation of a NP like “*fumeur*” should be along the lines of the one in (18).

$$(18) \quad [[fumeur]] = \lambda x_e \cdot x \text{ usually/habitually smokes.}$$

It seems to me that describing a smoker by saying that he or she “habitually” smokes is not precise enough. For example, if someone smokes every 1st of January, it seems

habitual for this person to do so, but it does not mean that she/he is a smoker. I will thus build on W&Z's account for my own analysis, but I will attempt to give a more precise semantics to the deverbal noun “*fumeur*”, and I will provide a more articulated analysis for the operator, corresponding to the generic operator used by Larson and W&Z.

#### 4. ANALYSIS

In this analysis, I will follow Winter and Zwarts (2012a,b), who argue that a decompositional analysis of deverbal nouns is necessary. Deverbal nouns contain a verbal root that denotes a set of events and an operator,  $EUR$ , whose semantic contribution is to take a predicate of events and give back a predicate of individuals, specifying the agent. This operator saturates the event argument contained in the verbal root. There are two possibilities for accounting for the fact that the positions of the operator  $EUR$  and the morphological *-eur* (semantically vacuous) are distinct:

1. Either there are two different positions, and the morpheme *-eur* is licensed by the presence of a covert operator,  $EUR$ .
2. Or there is only one element that appears higher than the verbal root at LF, and a spell-out rule specifies that this element must attach itself with the verbal root at PF.

In the semantics of the  $EUR$  -operator, I will make use of the concept of *situations* (see Kratzer 2004). A person will qualify as a smoker if he or she has a high chance of lighting a cigar/cigarette/... when he or she is allowed to, is able to, and wants to smoke. Thus, I propose that we can quantify over situations to obtain this kind of meaning. The denotation of a NP like “*fumeur*” is as in (19), and that of the  $EUR$ -operator is in (20).

- (19)  $[[\text{smoker}]] = \lambda x_e. \text{MOST}(s) [s \text{ is a situation in which the conditions for } x \text{ to smoke are good, } \exists E[E \in s \wedge \text{smoke}(E) \wedge \text{ag}(E) = x]]$
- (20)  $[[EUR]] = \lambda P_{\langle v, t \rangle}. \lambda x_e. \text{MOST}(s) [s \text{ is a situation in which the conditions for } x \text{ to } P \text{ are good, } \exists E[E \in s \wedge P(E) \wedge \text{ag}(E) = x]]$

Concerning the semantics of the adjective “*gros*” under its non-intersective reading, I will adopt the *i*-sum operator defined by Link (1983). I claim that the denotation for the non-intersective reading should modify the event argument present in the verbal root, taking a two-place predicate of simple atomic events (of type  $\langle v, t \rangle$ ), and giving back a predicate of complex events themselves constituted of atomic events (also of type  $\langle v, t \rangle$ ). The denotation for the non-intersective version of “*gros*” is given below.

- (21)  $[[\text{gros}_1]] = \lambda P_{\langle v, t \rangle}. \lambda E_v. E \text{ is constructed via } \oplus \text{ from many atomic events that are non-overlapping in } P(E)$

With this kind of semantic value for the non-intersective “*gros*”, a DP like “*un gros fumeur*” will denote someone who participates in many events of smoking when the

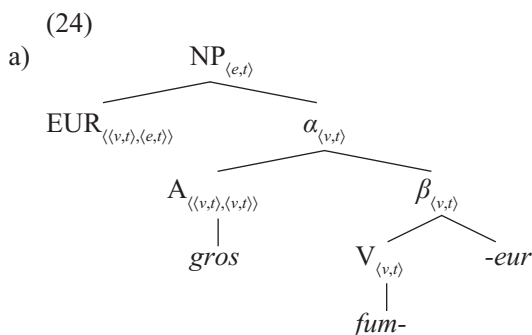
conditions for him or her are good to do so. This pairs well with the fact that a sentence containing the adverb “*beaucoup*” conveys the same meaning as one containing the adjective and the noun (examples (12) and (13) from section 1 are repeated below as (22) and (23)).

- |      |   |      |   |
|------|---|------|---|
| (22) | <i>C'est un gros fumeur.</i><br>it-is a GROS smoker<br>“He’s a heavy smoker.” | (23) | <i>Il fume beaucoup.</i><br>he smokes lot<br>“He smokes a lot.” |
|------|---|------|---|

#### 4.1 Prenominal Position: Non-Intersective Reading

Now that all the ingredients necessary for my analysis are introduced, let us consider how to derive the non-intersective interpretation of a DP containing a deverbal noun and the adjective “*gros*” in prenominal position.

The adjective, being of type  $\langle v,t \rangle$ , will apply to the verbal root *fum-* before the EUR-operator is inserted into the structure. The structure and the semantic composition of “*gros fumeur*” are shown below.



- b)  $[[\beta]] = \lambda_{E_v}. \mathbf{smoke}(E)$   
 $[[A]] = \lambda P_{\langle v,t \rangle}. \lambda_{E_v}. E$  is constructed via  $\oplus$  from many atomic events that are non-overlapping in  $P(E)$   
 $[[\alpha]] = \lambda_{E_v}. E$  is constructed via  $\oplus$  from many atomic events that are non-overlapping in  $\{E: \mathbf{smoke}(E)\}$   
 $[[NP]] = \lambda x_c. \text{MOST}(s)$  [ $s$  is a situation in which the conditions for  $x$  to **smoke** are good,  $\exists E[E \in s \wedge E$  is constructed via  $\oplus$  from many atomic events that are non-overlapping in  $\{E: \mathbf{smoke}(E)\} \wedge \mathbf{ag}(E) = x$ ]]

#### 4.2 Prenominal and Postnominal Position: Intersective Reading

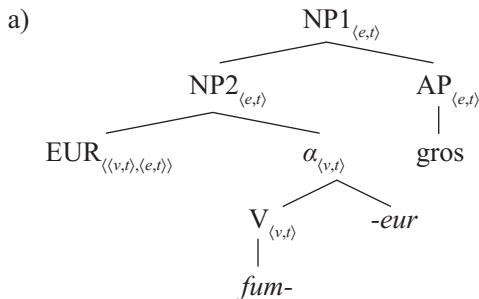
Concerning the postnominal position of the adjective “*gros*”, which only gives rise to an intersective reading, and the prenominal position, where it is also possible (even if

dispreferred) to have an intersective reading, I propose that the adjective has a different semantic value. This second version of the adjective is simpler, since it only denotes an individual property, as shown in (25).

$$(25) \quad [[\text{gros}_2]] = \lambda x_e. x \text{ is fat}$$

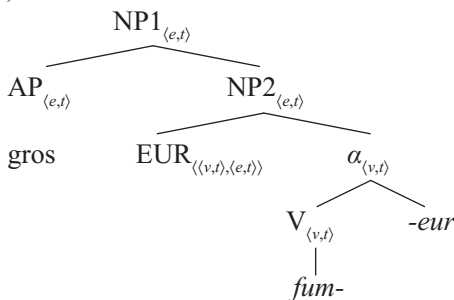
This adjective cannot modify a predicate of events such as the verbal root *fum-*, since it does not bear any event argument in its semantic value. The only option for applying it is after the combination of the verbal root and the EUR-operator. This gives rise to the intersective reading, as expected (consider (26) for the postnominal position, and (27) for the prenominal position with an intersective reading<sup>4</sup>).

(26)



- b)  $[[\text{NP2}]] = \lambda x_e. \text{MOST}(s) [s \text{ is a situation in which the conditions for } x \text{ to } \mathbf{smoke} \text{ are good, } \exists E[E \in s \wedge \mathbf{smoke}(E) \wedge \mathbf{ag}(E) = x]]$   
 $[[\text{AP}]] = \lambda x_e. \mathbf{fat}(x)$   
 $[[\text{NP1}]] = \lambda x_e. \text{MOST}(s) [s \text{ is a situation in which the conditions for } x \text{ to } \mathbf{smoke} \text{ are good, } \exists E[E \in s \wedge \mathbf{smoke}(E) \wedge \mathbf{ag}(E) = x] \wedge \mathbf{fat}(x)]$

(27)



4 For (27), the details of the semantic composition are not given, since they are the same as the ones for (26).

At this point, we must find a way to prevent a structure like the one in (28), where the adjective is postnominal and at the same time within the scope of the operator.

(28) [<sub>NP</sub> EUR [ [ [<sub>v</sub> *fum-* ] [ *-eur* ] ] *gros* ]]

In order to avoid these structures, I postulate the existence of a reduced relative clause in the case of postnominal adjectives, in a manner similar to that of (Cinque 2010):

(29) Un fumeur (*qui est*) gros  
 a smoker (*who is*) fat

In this way, the postnominal adjective can never be within the scope of the operator, and is not required to have access to the event argument inside the verbal root. Contrary to previous analyses in which the ambiguity of a DP like *a beautiful dancer* or “*un gros fumeur*” was either due to the complexity of the adjective (*A-analysis*, or “blame AP”) or that of the noun (*N-analysis*, or “blame NP”), the present proposal claims that this ambiguity is due both to the semantics of the deverbal noun and the presence of two lexical entries for the relevant ambiguous adjective (“blame both”). This claim supports the fact that an adjective in the same category as “*gros*” – one that has actually two different meanings – cannot be used ambiguously with every type of noun. As an example, in (30), the noun “*porte*” (“door”) does not bear any event structure. Thus, the adjective “*gros*” can only be used intersectively with this noun.

(30) Une grosse porte  
 a GROS.fem door

This demonstrates that the ambiguity of the adjective is not sufficient to create an ambiguity at the NP-level, and that a “blame both” type of analysis is on the right track.<sup>5</sup>

### 4.3 Bracketing Paradoxes

As W&Z (see 2012b: 642) emphasise, the kind of analysis proposed here raises the problem of what has been referred to as “bracketing paradoxes” in the literature. Namely, we need the adjective “*gros*” to apply to the root “*fum-*” before the semantically covert EUR applies. It should be useful to consider the analysis of *rebracketing* (Spencer 1988) to avoid this kind of problem (see examples (31) and (32)).

---

5 It is nonetheless possible that other nouns that are not -eur deverbal nouns can also combine with intersective and non-intersective versions of the adjective “*gros*”. In this case, we could imagine that this kind of noun also has some kind of complex structure (even if it is not morphologically visible), and that the two versions of the adjective modify the noun at different positions in this structure.



- (31) a. *Uneasier*: [un- [easi- -er]]  $\Rightarrow$  [[un- easi-] -er]  
 b. *Nuclear physicist*: [nuclear [physics- -ist]]  $\Rightarrow$  [[nuclear physics-] -ist]  
 (32) *Gros fumeur*: [gros [fum- -eur]]  $\Rightarrow$  [[gros fum-] -eur]

#### 4.4 Unifying the Two Lexical Entries for “*gros*”

In the analysis defended here (“blame both”), two different lexical entries for the adjective “*gros*” are required. However, intuitively, it seems a good idea to attempt to unify the two readings of the adjective. To do so, we can imagine an alternative semantics.

In this analysis, the type of the adjective is  $\langle\langle\sigma, t\rangle, \langle\sigma, t\rangle\rangle$ , where  $\sigma$  denotes the union of the domain of entities and the domain of events:  $\sigma = D_e \cup D_v$

If we consider that there is a type  $\sigma$  that can denote both events and individuals, we can make use of *partial functions* to avoid problems:

- (33) a.  $\lambda x_\sigma : x$  is an event.  $x \dots$   
 b.  $\lambda x_\sigma : x$  is an individual.  $x \dots$

In this case, the two versions of “*gros*” would have an identical type. Their meaning, however, is not related in any sense. To connect the two meanings, a possibility would be to do so inside the lexicon. This can be made possible in a generative lexicon framework, such as the one developed in Pustejovsky (1995). Since developing an account combining my analysis and the framework developed by Pustejovsky would go beyond the scope of this paper, I will leave this for further research (however, see also McNally (2005) for a similar argument).

## 5. CONCLUSION

In this study, I aim to give an account of adjectival ambiguity with deverbal nouns in French. I claim that the adjective “*gros*” is ambiguous between two different readings, one intersective and the other non-intersective. However, the ambiguity of these constructions is also due to the presence of an operator inside deverbal nouns. The scope of this operator defines where the intersective and non-intersective adjectives can apply.

An interesting development of the analysis developed here has to do with episodicity and genericity. It seems that the non-intersective reading of a complex NP containing the pronominal adjective *gros* and an *-eur* deverbal noun is correlated with the possibility of having an episodic reading of the deverbal noun (vs. a generic one). For someone to be “*un fumeur*”, this person must have been involved in episodic events of smoking.

Another kind of deverbal nouns, namely, instrumentals, do not follow the same logic. “*Un aspirateur*” (*a vacuum-cleaner*) is an object the purpose of which is to “*aspirer*”, but even if it has never been involved in an episodic event of “*aspirer*”, it is nonetheless referred to as “*un aspirateur*”.

What is interesting is that contrary to episodic deverbal nouns such as “*fumeur*”, non-episodic deverbal nouns do not give rise to a non-intersective reading when combined with the adjective “*gros*”: “*un gros aspirateur*” is not something that “vacuums

up a lot”, but an object that is big (see Roy and Soare 2014 for more details). This would need to be accounted for in the type of analysis presented here. There should be a difference in the semantics of deverbal nouns such as “*fumeur*”, and instrumentals such as “*aspirateur*”, that concerns the contribution of the EUR operator.

## References

- BRANCA-ROSOFF, Sonia/Serge FLEURY/Florence LEFEUVRE/Mat PIRES (2012) *Discours sur la Ville. Présentation du Corpus de Français Parlé Parisien des Années 2000* (CFPP2000). <http://cfpp2000.univ-paris3.fr/CFPP2000.pdf>.
- CINQUE, Guglielmo (2010) *The Syntax of Adjectives: A Comparative Study*. Cambridge, MA: MIT Press.
- DESPIC, Miloje/Yael SHARVIT (2008) “Some “Non-Intersective” Adjectives are Truly Non-Intersective.” In: S. Lima/K. Mullin/B. Smith (eds), *Proceedings of North Eastern Linguistic Society 39*. Amherst: GLSA, 281–294.
- ESHKOL -TARAVELLA, Iris/Olivier BAUDE/Denis MAUREL/Linda HRIBA/Celine DUGUA/Isabelle TELLIER (2011) *Un Grand Corpus Oral “Disponible”: Le Corpus d’Orléans 1968-2012*. <http://eslo.huma-num.fr/>.
- KAMP, Hans/Barbara Hall PARTEE (1995) “Prototype Theory and Compositionality.” *Cognition* 57/2, 129–191.
- KRATZER, Angelika (2004) “Covert Quantifier Restrictions in Natural Languages.” Lecture from Milan Meeting, Palazzo Feltrinelli (Gargnano).
- LARSON, Richard K. (1998) “Events and Modification in Nominals.” In: D. Strolovitch/L. Lawson (eds), *Proceedings from Semantics and Linguistic Theory (SALT) VIII*. Ithaca, NY: Cornell University, 145–168.
- LINK, Godehard (1983) “The Logical Analysis of Plurals and Mass Terms: A Lattice-Theoretical Approach.” In: R. Bäuerle/C. Schwarze/A. von Stechow (eds), *Meaning, Use and Interpretation of Language*. Berlin, Germany: De Gruyter.
- MCNALLY, Louise (2005) “Lexical Representation and Modification within the Noun Phrase.” *Recherches Linguistiques de Vincennes* 34, 191–206.
- MORZYCKI, Marcin (2013) “The Lexical Semantics of Adjectives: More Than Just Scales.” Ms., Michigan State University. [https://msu.edu/~morzycki/work/papers/chapter\\_not scales.pdf](https://msu.edu/~morzycki/work/papers/chapter_not scales.pdf)
- PARTEE, Barbara Hall (1995) “Lexical Semantics and Compositionality.” In: D. Osherson (ed.), *Invitation to Cognitive Science*. Cambridge, MA: MIT Press, 311–360.
- PARTEE, Barbara Hall (2010) “Privative Adjectives: Subsective Plus Coercion.” In: R. Bäuerle/U. Reyle/T. E. Zimmermann (eds), *Presuppositions and Discourse: Essays Offered to Hans Kamp*. Bingley, UK: Emerald Group Publishing, 273–285.
- PUSTEJOVSKY, James (1995) *The Generative Lexicon*. Cambridge, MA: MIT Press.
- ROY, Isabelle/Elena SOARE (2014) “On the Internal Eventive Properties of -er Nominals.” *Lingua* 141, 139–156.
- SIEGEL, Muffy Emily Ann (1976) *Capturing the Adjective*. Ph.D. Thesis. University of Massachusetts.

- SPENCER, Andrew (1988) "Bracketing Paradoxes and the English Lexicon." *Language* 64, 663–682.
- WINTER, Yoad/Joost ZWARTS (2012a) "Event Orientated Adnominals and Compositionality." Unpublished Ms., to appear in *Proceedings of Annual Meeting of the Israeli Association of Theoretical Linguistics*, October 2012, Tel-Aviv.
- WINTER, Yoad/Joost ZWARTS (2012b) "On the Event Semantics of Nominals and Adjectives – The One Argument Hypothesis." In: A. A. Guevara/A. Chernilovskaya/R. Nouwen (eds), *Proceedings of Sinn und Bedeutung 16*. Cambridge, MA: MIT Working Papers in Linguistics, 639–645.

#### Summary

### AMBIGUOUS ADJECTIVES IN FRENCH: THE CASE OF "GROS" WHEN COMBINED WITH DEVERBAL NOUNS

In this paper, I analysed the interaction between the French adjective "gros" and *-eur* deverbals. This adjective gives rise to a preferred non-intersective reading when it is in a prenominal position, but only an intersective reading when it appears after the noun. I claim that it is necessary to take into account both the semantics of the adjective and the semantics of the noun to account for the ambiguity present at the DP-level (a "blame both" analysis). An abstract operator, *EUR*, is always present within deverbals such as "fumeur" ("smoker"), and is partially responsible for the interpretation of the DP: if the adjective is within the scope of this operator, the DP will assume a non-intersective reading, whereas when it is outside of its scope, the DP bears an intersective interpretation. The adjective "gros" itself actually has two semantic values: one modifies the event argument present in deverbals, and the other modifies the individual argument of the noun (the agent of the verbal root).

**Keywords:** adjectives, ambiguity, deverbals, event semantics

#### Povzetek

### DVOUMNI PRIDEVNIKI V FRANCOŠČINI: PRIMER PRIDEVNIKA „VELIK“ V KOMBINACIJI Z IZGLAGOLSKIMI SAMOSTALNIKI

V članku je podana analiza interakcije med francoskim pridevnikom *gros* ‚velik‘ in izglagolskimi izpeljanimi samostalniki s pripono *-eur*. Pri tem pridevniku v položaju pred samostalniškim jedrom je najbolj pogosto nepresečno branje, medtem ko je v položaju za samostalnikom edino možno branje presečno. V članku zagovarjam trditev, da je za razlago dvoumnosti na nivoju določilniške zveze (DZ) potrebno upoštevati semantične lastnosti obeh, pridevnika in samostalnika. Pri izglagolskih samostalniških izpeljankah kot *fumeur* ‚kadilec‘ je vedno prisoten abstraktni operator *EUR*, ki je delno odgovoren za interpretacijo DZ. Če je pridevnik v dosegu operatorja, bo DZ imela nepresečno branje, če pa je pridevnik izven dosega, bo DZ imela presečno branje.

Pridevnik *gros* ‚velik‘ ima namreč dve semantični vrednosti: ena od njih modificira dogodek, ki je prisoten v izglagolskih samostalnikih, druga pa posamezni argument samostalnika (vršilec glagolskega dejanja v glagolskem korenu).

**Ključne besede:** pridevniki, dvoumnost, izglagolski samostalniki, dogodkovna semantika





## LEXICAL STRATA AND VOWEL (DIS)HARMONY: THE TURKISH TRANSFORMATION OF A BALKAN HYPOCORISTIC

### 1. INTRODUCTION

One shared feature of the Balkan sprachbund is a hypocoristic suffix with a vowel+palatal fricative ([oʃ]). This suffix is apparently Albanian in origin, possibly originating from an Albanian participial form in words such as *mekf* ‘calf’ and then extended (Camaj 1984; Berberi 1964). It is attested only with the vowel /u/ according to Newmark et al. (1982) but currently with both /u/ and /o/, (Bello, p.c.). A similar suffix appears in Romanian (/–uʃ/ and /–aʃ/, with additional variants for feminine forms). Polish and Serbian also share *osz*, as well as Hungarian.

Turkish also employs a version of this hypocoristic suffix. While it occasionally takes the form /–uʃ/ with the high back vowel, this variant is rare in Turkish. Rather, the default form is /–oʃ/, with the mid back rounded vowel. The high vowel variant would normally be subject to vowel harmony for backness and rounding, as shown in Table 1 below.

Table 1: Turkish high vowel harmony

Root vowels	Suffixal vowels
u, o	u
ü, ö	ü
e, i	i
a, ı	ı (î)

The mid back vowel, however, does not typically undergo vowel harmony. In fact, it does not typically appear in suffixes at all. The exception is a different, and far more common, hypocoristic suffix, which is simply the vowel /o/ on its own, suffixed to the initial (C)VC sequence of a name. According to Lewis (2000), /–oʃ/-suffixation is less familiar, but more socially acceptable, than simple /o/-suffixation.

The examples in (1) illustrate the application of /–oʃ/-suffixation in Turkish. It applies typically to the first syllable of the name (though this is not true of other Balkan

\* walter@metu.edu.tr

languages including the apparent source language, Albanian, in which substantially larger stems may be used as affixal bases, (Bello, p.c.)).

Within the maximal first syllable of the source and output forms, as many consonants are preserved as are phonotactically permissible as a coda in Turkish. In the output form, these then constitute the coda+second syllable onset, or simply onset in the case of a single consonant. This also distinguishes this suffixation process from simple /o/-suffixation, in which only one medial consonant is preserved.

Thus in Example (1a) below, only the initial consonant of the medial cluster is preserved, since /tm/ is not an acceptable coda cluster in Turkish, despite the fact that the /m/ is in any case resyllabified in the hypocoristic output as in the source name. However, in (1b), both medial /m/ and /b/ are preserved. (Examples are given in standard Turkish orthography, in which /ʃ/ is written as /ş/).

- |        |               |   |               |                 |
|--------|---------------|---|---------------|-----------------|
| (1) a) | <i>Fatma</i>  | > | <i>Fatoş</i>  |                 |
| b)     | <i>Pembe</i>  | > | <i>Pembiş</i> |                 |
| c)     | <i>bebek</i>  | > | <i>bebiş</i>  | ‘baby’          |
| d)     | <i>Ali</i>    | > | <i>Aliş</i>   | * <i>Aloş</i>   |
| e)     | <i>Minik?</i> | > | <i>Minnoş</i> | * <i>Minniş</i> |

As shown in examples (1b-d) above, Turkish has innovated a suffixal variant with a high front vowel, in addition to forms in /o/ and (rarely) /u/. These /i/-vowel forms do not appear in other Balkan languages with the comparable suffix.

This innovation allows for the preservation of typical Turkish vowel harmony for backness and rounding in at least some forms, as in (1b) and (1c).

However, note that in (1d), the front vowel variant of the suffix is used in spite of the back vowel appearing in the source name. *Aloş* would be the expected form, given the normal application of vowel harmony in high vowels for backness (and also given the absence of the Turkish mid/back unrounded high vowel ‘ı,’ which would normally be used after /a/ but virtually never appears for this suffix).

Conversely, the front vowel /i/ would be expected in (1e), harmonizing with the identical root vowel. However, /i/ is never used in the suffix with this form, which is an essentially lexicalized one referring to pet cats.

It is clear that the innovation of a front-vowel form of this suffix in Turkish has not led to consistency and predictability in its suffixation according to vowel harmony, despite this being the presumable reason for its introduction. Lewis (2000) claims that use of /-oʃ/ versus its high front vowel variant /-iʃ/ is actually “not reducible to a rule.” The remainder of this study will focus on patterns of variation in suffixal vowel choice in both naturally-occurring and elicited speech data. Section 2 contains the naturalistic data, while Section 3 discusses the elicited data. In Section 4, I summarize and discuss the findings as a whole and implications for phonological theory. Section 5 concludes the paper.

## 2. TURKISH HYPOCORISTIC /-OÇ/-SUFFIXATION IN NATURALISTIC SPEECH

Forms in this section are gathered from naturalistic contexts such as overheard conversations and internet chat, as well as informal querying of 2 Turkish informants, added to the forms given in Lewis (2000). The resulting mini-corpus contains 32 potentially suffixable personal names.

### 2.1 Productivity

Three very common personal names were rejected by informal informants 1 and 2 as candidates for /-oç/-suffixation. In addition, a small number of the attested forms are essentially lexicalized and therefore of doubtful evidence for suffix productivity.

(2) Lexicalized suffixal names

- a) *Fatma* > *Fatoş*
- b) *Minik?* > *Minnoş*

However, application to novel and foreign names indicate some degree of productivity.

(3) Novel suffixed names

- a) *Enabet* > *Eniş*
- b) *Marlboro* > *Malbuş*

### 2.2 Variation

Two of the names are attested with two alternate hypocoristic forms.

(4) Alternating suffixed names

- a) *Mehmet* > *Memoş, Memiş*
- b) *Fatma* > *Fatoş, Fatış* (the former is clearly dominant)

### 2.3 Vowel quality

The front high rounded vowel is attested in a single form (*Mülayım*>*Mülüş*). The same is true for the mid/back unrounded vowel ‘ı.’ The back high rounded vowel is attested in only two (*Ertuş, Malbuş*).

Table 2: Distribution of suffixal vowels in naturalistic speech

	null	variable	ü	ı	u	i	o	total
<b>Tokens</b>	3	2	1	1	2	11	13	33
<b># harmonic</b>	/	/	1	1	1	7	1	11/28



As Table 2 shows, the remainder of the forms are relatively equally split between /i/ and /o/ vowels. However, this balance does not result from their use according to the vowel harmony rule. Only about a third of the hypocoristic outputs are vowel-harmonic (11 out of 28).

Which ones? About half of the /i/-vowel forms appear in vowel-harmonic outputs. This is substantially less than one would expect for a suffix that has been introduced presumably in order to satisfy vowel harmony. Forms with /o/, on the other hand, are almost never harmonic. Thus the behavior of the two primary vowel variants is split, but not in such a way as to consistently satisfy vowel harmony.

Instead, the ‘new’ vowel /i/ satisfies it half the time, and the ‘old’ vowel /o/, never.

The following overheard exchange perfectly encapsulates the array of variation observed so far. It occurred outside a daycare center between one of the daycare teachers, and the mother of one of the students (Fikriye). As Fikriye’s mother approached the entrance of the daycare, Fikriye was hiding behind her legs.

(5) Suffixation in context

Mother: *Fikros* *gelmedi* *bugün!*  
Fikriye come-neg.past this-day  
‘Fikriye didn’t come today!’

Teacher: *Nerede* *Fikris?!*  
Where Fikriye  
‘Where is Fikriye?!’

Here both vowel allomorphs are used in quick succession in the same conversation, albeit by different speakers. Their coexistence excited no interest or comment on the part of the participants. Yet the data so far suggest that the choice between vowels is not random. It is clear that more systematic data is necessary on the usage of these allomorphs.

### 3. TURKISH HYPOCORISTIC /-OÇ/-SUFFIXATION IN ELICITED DATA

In an effort to more systematically explore hypocoristic suffixation in Turkish, survey data was collected from 14 native Turkish speakers who gave their informed consent for participation. These speakers were of traditional university student age, studying in an English-language-teaching program in an English-medium university (average age, early 20s). In an initial informal discussion, participants were provided with a couple of examples of names with the suffix in question, and asked if they knew ones like this. All speakers had at least this much basic familiarity with such forms.

They were then given a written survey (in English) which asked them to check how often they had *heard* such forms, on a four-point scale of never, a little bit, sometimes, and very often. The next question asked them how often they had *used* such forms.

Nine of the 14 speakers reported that they *sometimes* or *very often* hear and use this type of nickname. This group included 7 female and 2 male speakers. The data analyzed

in this section comes from the responses of those speakers to a survey on this suffixation process. The survey consisted of a list of 50 gender-balanced common Turkish personal names, semi-randomly scattered throughout the alphabet with respect to first letter.

Participants were asked to write the possible nicknames for each name, as follows: “If both *oş* and *iş* are possible, you can write both. If neither one sounds good, you can leave the blanks empty. If a different form sounds good too, you can add it in the ‘other’ column. Don’t spend a lot of time thinking about each name – just write your first ideas and move on.”

At the end of the survey, participants were also asked to add any other forms with *oş/iş* which they had used or heard.

### 3.1 Productivity

All participants chose not to provide output forms for at least some of the 50 prompt names.

Table 3: Number of null outputs per subject from 50

Participant	1	2	3	4	5	6	7	8	9	Average
Null/50	15	10	16	19	23	5	6	29	16	15%

These null responses ranged from 10% to over 50% of the prompt forms, averaging 15% for the group. On the other hand, roughly half (4) of the participants provided additional suffixed forms for names not included in the prompt list in the space provided for them to do so.

### 3.2 Variation

One participant gave 2 or, more often, 3 variant forms for each non-null output (different vowels; n=5 null, 41 alternating non-nulls). This participant’s data is excluded from subsequent analysis. Trisyllabic outputs were produced by 3 subjects for a handful of forms (2, 2, and 5 each), with no apparent predictive factor. With respect to medial consonants, participants generally follow the rule identified above:

- (6) Medial consonant rule: Both preserved if sonority contour falls (~acceptable Turkish syllable coda), otherwise first one only is preserved. Exceptions:
  - a) All 6 responses from Zehra > Zehruş
  - b) One token Bahruş from Bahar (other 3 preserve only /h/)
  - c) Variation in sibilant+liquid sequences:
    - i. Esra > Esroş (4 participants, 2 w/o /r/)
    - ii. Özlem > ÖzlVş (3 participants, 2 w/o /l/)
  - d) One token Ibrahim > Ibroş, one token Tolga > Togiş

As for the choice of suffixal vowel, participants produced multiple possible variants for up to a quarter of the prompt names for which they provided outputs.

Table 4: Number of prompt names with variant outputs per subject

Participant	1	2	3	4	5	7	8	9	average
tokens varying/50	2	8	17	6	12	0	0	1	6%
% of total outputs	5	14	25	14	23	0	0	3	10.5%

These varied *only* with respect to suffixal vowel, not other factors. In Section (3c), each variant is counted independently.

In sum, apart from a small amount of variation with respect to medial consonants, variation is observed primarily in the choice of suffixal vowel.

### 3.3 Vowel quality

One token was produced with the mid front rounded vowel /ö/, none with ‘ı,’ and none with /ü/. Suffixes with /u/ are also marginal, with only 12 total outputs produced. This is ~10% or less of the number of tokens with either /i/ or /o/, respectively.

Table 5: Distribution of suffixal vowels in elicited forms

	i	o
tokens	114	188
% harmonic	46	35

As in the naturalistic data, /i/ and /o/ are the typical suffixal vowels for these allomorphs. In elicited data, the /o/ allomorph outnumbers the /i/ allomorph by approximately 50%, rather than being used at a comparable rate as seen in Section 2.

The elicited data parallels the naturalistic data in that disharmony is prevalent among suffixed forms, despite the introduction of a potentially harmonizing suffix. In addition, the same pattern is observed, in which suffixes with innovated /i/ are more likely to be harmonic compared to /o/ forms, but still violate vowel harmony more than half of the time.

### 3.4 Potential Conditioning Factors

One participant shared the intuition that /i/ suffixes were used with female names, and /o/ suffixes with male ones. This intuition was not borne out by the data; suffixal vowels are used in roughly the same proportions regardless of name gender (unisex names omitted).

Table 6: Distribution of suffixal vowels according to gender of name

	<b>i</b>	<b>o</b>
<b>feminine</b>	70	113
<b>masculine</b>	36	77

Place of articulation of the preceding consonant also appears to play no role in suffix selection.

Table 7: Distribution of suffixal vowels according to preceding consonant

	<b>i</b>	<b>o</b>
<b>alveolar</b>	56	127
<b>labial</b>	22	40
<b>palatal</b>	13	13
<b>velar</b>	11	10
<b>laryngeal</b>	7	5

As seen elsewhere, /o/ suffixes outnumber /i/ suffixes by approximately 2-to-1 for the two largest categories (alveolar and labial), regardless of consonant place. The numbers are roughly equivalent for other places of articulation (palatal, velar, laryngeal), but the smaller numbers of tokens involved means that this is more likely to happen by chance. It is clearly not the case that labial is conditioning a following rounded vowel, or palatal a following high front vowel, as might have been suspected.

Table 8 summarizes the likelihood of occurrence of the two suffixal vowel variants with different preceding vowels, grouped into backness-based classes.

Table 8: Distribution of suffixal vowels according to preceding vowel

<b>preceding vowel</b>		<b>i</b>	<b>o</b>
<b>FRONT</b>	E	41	95
	I	6	23
	Ö	8	5
	Ü	1	8
<b>TOTAL</b>		<b>56</b>	<b>131</b>
<b>BACK</b>	A	28	51
	U	14	11
	O	10	2
<b>TOTAL</b>		<b>52</b>	<b>64</b>

When the preceding stem vowel is a front vowel, the suffixal vowel is overwhelmingly likely to be a back vowel /o/, rather than a harmonizing front vowel. When the preceding stem vowel is a back vowel, the suffixal vowel is about equally likely to be either a front vowel or a back vowel.

Looking simply at the suffixal vowels themselves, we see that /i/ is about equally likely to be used after either a front vowel or a back vowel. On the other hand, /o/ is used more than twice as often after front vowels as after back vowels, yielding a majority of disharmonic outputs.

The case of /o/ forms after /u/ vowels presents an additional complication. Vowel sequences of /u-o/ are disallowed in Turkish, and speakers are highly aware of this. Nonetheless, such sequences in hypocoristic suffixed forms are almost equally likely as those with /i/, and much more likely than totally harmonic /o-o/. Here once again, hypocoristic suffixation creates gratuitous and non-faithfulness-motivated phonotactic violations.

#### 4. SUMMARY AND DISCUSSION

In both datasets, suffixal vowels follow similar patterns of distribution and disharmony. ‘Original’ /o/ suffixes are most prevalent, compared to innovative /i/ suffixes. They are also more likely to violate vowel harmony, compared to /i/ forms. However, even /i/ forms violate vowel harmony roughly half of the time. Why are such phonotactic violations introduced, in the absence of faithfulness considerations for foreign source forms?

##### 4.1 Disharmony Elsewhere in Turkish

Disharmonic forms in Turkish also arise from a number of other sources: both in forms borrowed from other languages, and in normally harmonizing suffixes when applied to roots known to be foreign, in some cases.

- (7) Disharmonic suffixes in loans
- |    |             |                           |                                     |
|----|-------------|---------------------------|-------------------------------------|
| a) | <i>harf</i> | ‘alphabet letter; Arabic’ | → <i>harfler</i> , * <i>harflar</i> |
| b) | <i>saat</i> | ‘hour, clock; Arabic’     | → <i>saatler</i> , * <i>saatlar</i> |

Vowel epenthesis in loanword onset clusters also tends to be disharmonic. This occurs even though like the hypocoristic suffix vowel, the onset epenthetic vowels are not constrained by faithfulness to any underlying vowel quality. Also like the hypocoristic forms, the epenthetic outputs are numerically dominated by a single vowel (/i/ for epenthesis, /o/ for hypocoristic forms). In both cases, this ‘default’ vowel is disharmonic a majority of the time. In both cases, the alternating form (/i/ for epenthesis, /i/ for hypocoristic forms) is also disharmonic a substantial portion of the time, (Walter 2014).

In neither case did alternation ‘solve’ the disharmony ‘problem.’ Instead, there is an increase in harmony violations overall for these two specific lexical strata: foreign-origin vocabulary and nicknames.

## 4.2 Nicknames as a Lexical Stratum

In previous work I refer to nicknames, along with given names, as examples of ‘paralinguistic speech’ – forms which inherently carry a connotation of social closeness, and also display specific phonological properties, (Walter 2009 and references therein).

This line of research reveals that names and nicknames are likely to be more marked/complex in form. English personal names preferentially include phonemes which have less frequent type frequency in the lexicon as a whole. That is, rarer segments such as /dʒ/ and /v/ are much more common in names than in other lexical items, whereas more frequent phonemes such /s/ and /t/ are underrepresented.

English is not alone in this respect. I document the same inverse relationship in personal names from Arabic, Greek, and Czech (Walter 2009). To take just one more example, Moroccan Arabic babytalk, also a sociolinguistically marked register indicating social closeness, preferentially includes marked segments such as pharyngeal and ‘emphatic’ pharyngealized consonants, as well as the phoneme /p/ which otherwise does not exist in Arabic (Ferguson 1982).

Members of this proposed lexical stratum also tend to undergo processes which increase phonological complexity. Although consonant repetition is cross-linguistically avoided generally (Frisch, Pierrehumbert and Broe 2004), such reduplicative structures are cross-linguistically common in nicknames. Nicknames undergo segmental processes such as fortition/affrication which are otherwise unattested, e.g. palatal affrication from /s/ to /tʃ/ in Spanish. Such fortition is observed across the board in nicknames such as Rosario>Charo and Concepción>Concha, yet is unattested as a phonological rule in any other language in the PBase database of phonological rules (Mielke 2008). A similar pattern is observed in ‘intensity shifts’ towards affrication in diminutive forms of West North American languages, (Nichols 1971). Finally, more complex syllable structures are often tolerated in such registers/strata.

In sum, then, nicknames and loanwords share the property of being more phonologically marked overall than the lexicon as a whole – of being *marked* by *markedness*.

## 4.3 Theoretical Approaches

This markedness effect is reminiscent of Ito and Mester’s well-known work on stratification in the Japanese lexicon (2009). The classic example is the application of the Japanese rule of *rendaku*. In *rendaku*, intervocalic voicing applies at compound boundaries in the native lexicon. However, there is no intervocalic voicing for borrowed forms. Thus, for loanwords only, faithfulness to underlying forms is maintained at the expense of markedness. Ito and Mester document a nested lexicon of native/unmarked forms with additional layers of foreign/marked forms.

However, this is *not* what we see in Turkish. To further analogize to *rendaku*, the Turkish /-of/ pattern (as well as loanword vowel epenthesis documented elsewhere), it is as if underlyingly voiced phonemes in borrowed forms are being gratuitously devoiced in Japanese – because Japanese speakers have noticed that unvoiced phonemes anomalously appear in foreign forms, and are carrying that generalization further by creating such violations even in violation of faithfulness constraints.

Ito and Mester's approach, in which more marked structures are allowed in foreign layers of the lexicon due to higher-ranking faithfulness constraints, cannot accommodate the Turkish data, in which the markedness observed is *not* due to faithfulness to foreign phonotactic violations. Rather, indexed constraints *enforcing* marked structures in such strata are required, such as those proposed by Pater (2010). In some sense, this exceptionality/markedness is purposeful, in that it serves to distinguish a separate stratum in the lexicon.

## 5. CONCLUSIONS

This particular Turkish hypocoristic pattern is best described as semi-productive. It is reminiscent of English suffixal rules such as velar softening, which may be productively applied but only to roots which they seem to 'fit' (i.e. Latinate-seeming; Pierrehumbert 2006). This marginality corresponds to its origin as a Balkan form which has spread only partially through the wider Turkish speech community. While considerable variability is observed for vowel quality in these suffixes, it is noteworthy that overall, disharmony/markedness is increased in this part of the lexicon – as seen in other languages' nicknaming patterns.

This increase is *not* due to properties of the source language, thus paralleling epenthesis in Turkish loanwords. The phenomenon is better modeled by process-specific constraints than faithfulness rerankings, in this case achieving sociolinguistic goals.

## References

- BERBERI, Dilaver (1964) *Phonological and morphological adaptation of Turkish loanwords in contemporary Albanian Geg dialects of Kruja: A synchronic analysis*. Ph.D. thesis, Indiana University.
- CAMAJ, Martin (1984) *Albanian grammar*. Wiesbaden: Otto Harrassowitz Verlag.
- FERGUSON, Charles (1982) "Baby-talk in six languages." *American Anthropologist* 66/6, 103–114.
- FRISCH, Stefan/Janet PIERREHUMBERT/Michael BROE (2004) "Similarity avoidance and the OCP." *Natural Language and Linguistic Theory* 22, 179–228.
- ITO, Junko/Armin MESTER (2009) "Lexical classes in phonology." In: S. Miyagawa/M. Saito, (eds), *Handbook of Japanese linguistics*. Oxford: Oxford University Press, 84–106.
- LEWIS, Geoffrey (2000) *Turkish grammar*. 2<sup>nd</sup> edition. Oxford: Oxford University Press.
- MIELKE, Jeff (2008) *The emergence of distinctive features*. Oxford: Oxford University Press.
- NEWMARK, Leonard/Philip HUBBARD/Peter PRIFTI (1982) *Standard Albanian: A reference grammar for students*. Stanford: Stanford University Press.
- NICHOLS, Johanna (1971) "Diminutive consonant symbolism in western North America." *Language* 47, 826–848.

- PATER, Joe (2010) "Morpheme-specific phonology: Constraint indexation and inconsistency resolution." In: S. Parker (ed.), *Phonological argumentation: Essays on evidence and motivation*. London: Equinox, 123–154.
- PIERREHUMBERT, Janet (2006) "The statistical basis of an unnatural alternation." In: L. Goldstein/D. H. Whalen/C. Best (eds), *Laboratory Phonology 8, Varieties of phonological competence*. Berlin: Mouton de Gruyter, 81–107.
- WALTER, Mary Ann (2009) *Sound patterns in paralinguistic speech*. Colloquium presentation at Northwestern University, Evanston, IL, USA.
- WALTER, Mary Ann (2014) *The significance of vowel insertion in Ottoman Turkish*. Paper presented at the 17<sup>th</sup> International Conference on Turkish Linguistics, Rouen, France.

### Summary

#### LEXICAL STRATA AND VOWEL (DIS)HARMONY: THE TURKISH TRANSFORMATION OF A BALKAN HYPOCORISTIC

In this study I explore the phonological behavior of the hypocoristic suffix /-oʃ/-iʃ/ in Turkish. Such a suffix is common to many of the Balkan languages. Turkish differs in its introduction of the front vowel variant of the suffix, presumably to satisfy the vowel harmony requirements in Turkish for backness and rounding in high vowels. However, in spite of the potentially alternating suffix allomorphs, collection of naturalistic data as well as of elicited survey data reveals that the majority of nickname outputs are disharmonic. I conclude that the Turkish data provides further evidence for Ito and Mester's (2009) key insight that different strata of the lexicon may operate according to different rules/constraint rankings.

However, the Turkish data is not consistent with their specific faithfulness-based approach. The hypocoristic lexical stratum exhibits a greater number of vowel harmony violations, but *not* due to more faithfulness to vowel inputs/underlying forms. Rather, the harmony violations in this stratum are gratuitous – I argue, precisely in order to distinguish this stratum from the lexicon at large. An approach such as Pater's (2010) indexed constraints model better accommodates this type of lexical variation.

**Keywords:** Balkan, hypocoristics, loanwords, phonology, Turkish

### Povzetek

#### LEKSIKALNI SLOJI IN SAMOGLASNIŠKA (DIS)HARMONIJA: SPREMEMBA BALKANSKE HIPOKORISTIKE V TURŠČINI

Delo proučuje, kako se s stališča glasoslovja obnaša turška pripona za tvorjenje hipokorističnih izrazov /-oʃ/-iʃ/, ki je sicer skupna več jezikom na Balkanu. Turščina se od ostalih jezikov razlikuje pri rabi različice s sprednjim samoglasnikom, in sicer domnevno zaradi upoštevanja samoglasniške harmonije v turščini, ki narekuje nazaj



pomaknjen izgovor in zaokroženost pri visokih samoglasnikih. Kljub potencialno spreminjajoči se priponski alomorfiji pa podatki, pridobljeni v raziskavi pri naravnih govorcih jezika v naravnem okolju in v kontroliranem poskusu, kažejo na to, da večina ljubkovalnih imen ne podlega samoglasniški harmoniji. Iz analize zaključimo, da so turški podatki še en dokaz, ki potrjuje ključno ugotovitev v Ito/Mester (2009), da različni sloji v leksikonu lahko delujejo po različnih pravilih oziroma imajo različno rangiranje omejitev.

Turški podatki pa niso skladni s specifičnim okvirom v Ito/Mester (2009), ki je osnovan na zvestobnostnih omejitvah. Sloj hipokorističnih besed namreč kaže večje število kršitev samoglasniške harmonije, vendar ne zaradi zvestobe samoglasniškemu vnosu/globinski obliki, ampak se kršenje harmonije pojavi prav zato, da se ta sloj razlikuje od leksikona na splošno. Ta tip leksikalne variacije bolje razloži model indeksiranih omejitev v Pater (2010).

**Ključne besede:** Balkan, hipokoristiki, prevzete besede, fonologija, turščina



## SOME ASPECTS OF AGREEMENT WITH NUMERAL PHRASES IN POLISH\*\*\*

### 1. INTRODUCTION

Patterns of agreement between a quantified subject and a verbal or adjectival predicate/participle in Polish seem to pose a challenge to standard theories of Agree (cf. Chomsky 2000, 2001, 2008) as we observe a full subject-verb agreement, i.e., in person, number and gender, with subjects quantified by lower numerals, i.e., <5, e.g., (1a), whereas phrases with higher numerals, i.e., ≥5, and numeral quantifiers exceptionally force default agreement, i.e., 3<sub>SG,NEUT</sub>, e.g., (1b).<sup>1</sup>

- (1) a) *Trzy kobiety weszły do budynku.*  
three<sub>NOM</sub> women<sub>NOM,FEM,PL</sub> entered<sub>FEM,PL</sub> to building  
'Three women entered the building.'

\* wjacek@wa.amu.edu.pl

\*\* dziubala@wa.amu.edu.pl

\*\*\* Extensive parts of this material were presented at FDSL 10.5 in 2014 (4-6 December, Brno), The Agreement Across Borders Conference in 2015 (15-16 June, Zadar) and Slavic Linguistics Society in 2015 (4-5 September, Heidelberg), and appeared in Witkoś and Dziubała-Szrejbrowska (2016). This contribution is partially funded by grant no. 2012/07/B/HS2/02308 issued by the Polish Science Center.

1 Virile (masculine personal) lower numerals also trigger default agreement, i.e., 3<sub>SG,NEUT</sub>, in Polish, e.g.

- (i) *Dwóch/pięciu mężczyzn wyszło z budynku.*  
two/five<sub>VIR</sub> men<sub>GEN,VIR,PL</sub> left<sub>3SG,NEUT</sub> from building  
'Two men left the building.'

Yet, due to the syncretism between genitive and accusative in virile plural the case optionality in agreement with predicative adjectives and participles is not detectable, e.g.

- (ii) *Pięciu mężczyzn zostało wybranych do rady wydziału.*  
five<sub>ACC</sub> men<sub>GEN</sub> was<sub>3SG,NEUT</sub> selected<sub>ACC/GEN</sub> to council faculty  
'Five men were selected for the faculty council.'

- (iii) Tom spotkał **pięciu mężczyzn.**  
Tom met five<sub>ACC</sub> men<sub>ACC</sub>.  
'Tom met five men.'

- (iv) Tom nie widział **pięciu mężczyzn.**  
Tom not see five<sub>GEN</sub> men<sub>GEN</sub>.  
'Tom did not see five men.'

- b) *(tamte) siedem/wiele kobiet weszło do budynku.*  
 those<sub>ACC.PL</sub> seven<sub>ACC</sub>/many women<sub>GEN.FEM.PL</sub> entered<sub>3SG.NEUT</sub> to building  
 ‘Those seven/many women entered the building.’

Yet another puzzle in accounting for agreement facts emerges in the context of adjectival predicates and participles found with quantified subjects (QPs) as the predicate/participle may optionally occur in accusative or genitive, which indicates agreement either with the numeral or the nominal, e.g., (2).<sup>2</sup>

- (2) *Pięć kobiet było wybrane/wybranych do rady nadzorczej.*  
 five<sub>ACC</sub> women<sub>GEN.PL</sub> was<sub>3SG.NEUT</sub> chosen<sub>3PL.ACC/3PL.GEN</sub> for board supervisory  
 ‘Five women were chosen for the supervisory board.’

Interestingly, in other related languages, for instance in Russian, agreement with quantified subjects is optional, e.g., (3a); however, once at least one element of the phrase is nominative, then only a full agreement becomes felicitous, e.g., (3b).

- (3) a) *Pjat’ devušek rabotali/rabotalo tam.* (Russian)  
 five girls<sub>GEN.PL</sub> worked<sub>PL/SG</sub> there  
 ‘Five girls worked there.’  
 b) *Ėti pjat’ devušek rabotali/\*rabotalo tam.*  
 these<sub>NOM</sub> five girls<sub>GEN.PL</sub> worked<sub>PL/SG</sub> there  
 ‘These five girls worked there.’

The peculiarities of agreement patterns in these languages have led us to resume a discussion of different facets of agreement, but this time, in the light of a nanosyntactic approach to grammar (cf. Starke 2009; Caha 2009, 2010 inter alia) which seems to adequately capture problematic paradigms with GoQ.

## 2. THE AGREEMENT PUZZLE

In numerous attempts to account for agreement patterns it has been proposed that the form of the predicate, i.e., agreeing with the subject or default, i.e., 3<sub>SG.NEUT</sub>, depends on the structure of the quantified subject which can constitute either a DP or QP (e.g., Pesetsky 1982; Franks 1994, 1995; Pereltsvaig 2006).<sup>3</sup> In another approach, Bošković (2006) submits that a division into QP/DP can be dispensed with in favor of

2 In oblique case positions, higher numerals, i.e.,  $\geq 5$ , behave like typical adjectival modifiers, i.e., they agree in number gender and case with the modified noun, e.g.

*Nauczyciel wyszedł z siedmioma uczniami.*  
 teacher left with seven<sub>INST</sub> pupils<sub>INST</sub>.  
 ‘A teacher left with seven pupils.’

3 For a detailed discussion of DP/QP status of quantified phrases in Polish see Witkoś and Dziubała-Szrejbrowska (2015).

a distinction based on (abstract) case. He puts forward the idea that high numerals are ambiguous between case (nominative/accusative) and caseless forms, whereas nominative always correlates with a full agreement. Moreover, the numeral ‘pjat’ occupies the specifier position of a nominal functional projection FP and shows case properties in one syncretic form:<sup>4</sup>

- (4) [<sub>FP</sub> QP [<sub>F</sub> F NP]]  
 (5) pjat’: a. nominative b. accusative c. caseless

According to Bošković (2006) nominative case marking entails full agreement (as in Chomsky 1995 and the T/Agr<sub>s</sub> model). Moreover, Bošković provides example (3), repeated here as example (6), arguing that as soon as any element of the FP shows nominative the entire FP loses its ambiguity, it transpires as nominative and it triggers full agreement:

- (6) a) *Pjat’ devušek rabotali/rabotalo tam.* (Russian)  
 five girls<sub>GEN.PL</sub> worked<sub>PL/SG</sub> there  
 ‘Five girls worked there.’  
 b) *Ėti pjat’ devušek rabotali/\*rabotalo tam.*  
 these<sub>NOM</sub> five girls<sub>GEN.PL</sub> worked<sub>PL/SG</sub> there  
 ‘These five girls worked there.’

A corresponding pattern, however, is not found in Polish as none of the elements of the quantified phrase is nominative (the numeral is marked with ACC, the noun with GEN and the determiner with ACC or GEN). Thus, as predicted in (5), we always observe a default agreement in Polish, e.g.:

- (7) a) *Pięć dziewczyn \*pracowały/pracowało tam.* (Polish)  
 five<sub>ACC</sub> girls<sub>GEN.PL</sub> worked<sub>PL/SG</sub> there  
 ‘These five girls worked there.’  
 b) *Te/tych pięć dziewczyn \*pracowały/pracowało tam.*  
 these<sub>ACC, GEN.PL</sub> five<sub>ACC</sub> girls<sub>GEN.PL</sub> worked<sub>PL/SG</sub> there  
 ‘These five girls worked there.’

Considering (5), we propose to credit the difference between Polish and Russian to distinct cases on the QP subject. While in Russian the case of the high numeral (and certain quantifiers) varies between nominative and accusative, in Polish it is accusative. So Russian T can successfully probe for the  $\phi$ -features of the subject QP when it shows the  $\phi$ -features that would match T. In Polish, and in certain contexts in Russian, the high numeral appears in the subject position in the other structural case, accusative

4 This proposal corresponds to *The Licensing Parameter* from Franks (2002): Polish QPs are licensed only in accusative DPs; Russian QPs are licensed in accusative and nominative DPs; SC QPs are licensed in all case DPs.

(cf. Franks 1994, 1995, 2002; Przepiórkowski 2004), which precludes agreement for  $\phi$ -features with T:<sup>5</sup>

$$(8) T_{\phi/default} \leftrightarrow [_{QP[+acc]} Q [ NP ]]$$

Following this imperfect Match, T defaults to  $3_{SG.NEUT}$ , which suffices to account for (7) above. However, we are still left with the issue of the optional participial and adjectival agreement:<sup>6</sup>

- (9) *Te/tych*            *pięć*    *dziewczyn*    *było*            [<sub>Prtp</sub> *wybran-e/-ych*  
 these<sub>ACC.PL/GEN.PL</sub>    five<sub>ACC</sub>    girls<sub>GEN.PL</sub>    was<sub>3SG.NEUT</sub>    selected<sub>3PL.ACC/3PL.GEN</sub>  
*do konkursu*].  
 to contest  
 ‘These five girls were selected for the contest.’

### 3. GENITIVE OF QUANTIFICATION: A BRIEF DIACHRONIC DISCUSSION

An adequate analysis of agreement patterns with quantified subjects requires a proper understanding of the nature of numerals and changes they underwent. Considering that genitive is a typical adnominal case as well as the fact that in the past numerals  $\geq 5$  used to be nouns with a feminine declension, the question that should be initially addressed is what category numerals are and what feature make-up they possess.

Historically, Polish higher numerals shifted from pure nominal (feminine declension) forcing their complement to appear in genitive to functional elements serving as

5 We follow Przepiórkowski (1999, 2004), who proposes that QPs are marked for accusative in the subject position on the basis of the following comparison, among others:

- (i) *(tych/te)*                                    *pięć*                                    *kobiet*                                    *stało*.  
 these<sub>GEN.FEM</sub>//these<sub>NOM?/ACC.FEM</sub>    five<sub>NOM?/ACC.FEM</sub>                    women<sub>GEN.FEM.PL</sub>                    stood<sub>3SG.NEUT</sub>  
 ‘These five women were standing.’
- (ii) *(tych/\*ci)*                                    *pięciu*                                    *mężczyzn*                                    *stało*.  
 these<sub>ACC/GEN.MASC</sub>//these<sub>NOM.MASC?</sub>    five<sub>NOM?/ACC/GEN.MASC</sub>                    men<sub>GEN.MASC.PL</sub>                    stood<sub>3SG.NEUT</sub>  
 ‘These five men were standing.’

The common case form of the demonstrative (these) for both genders is accusative, on the assumption that its optional genitive reflects the placement of the demonstrative in the domain of the NP-complement and its subsequent raising to the domain of the numeral/quantifier.

6 The case of the demonstrative pre-quantifier, i.e., accusative or genitive, does not correlate with the case suffix on the adjective/participle, e.g.:

- (i) *Te*            *pięć*    *dziewczyn*    *zostało*    *wybranych*.  
 these<sub>ACC.PL</sub>    five<sub>ACC</sub>    girls<sub>GEN.PL</sub>    was<sub>3SG.NEUT</sub>    selected<sub>GEN.PL</sub>
- (ii) *Tych*            *pięć*    *dziewczyn*    *zostało*    *wybrane*.  
 these<sub>GEN.PL</sub>    five<sub>ACC</sub>    girls<sub>GEN.PL</sub>    was<sub>3SG.NEUT</sub>    selected<sub>ACC.PL</sub>  
 ‘These five girls were selected.’

For more examples see Przepiórkowski and Patejuk (2012). For a closer analysis of this case variability on the pre-quantifier, see Witkoś and Dziubała-Szrejnbrowska (2016).

modifiers agreeing in case with the nominal head. Rutkowski (2007: 240) presents the following comparison, e.g., *pięć lat* (five years).

(10)

Case	OLD POLISH	PRESENT DAY POLISH
NOMINATIVE	pięć lat	pięć lat
GENITIVE	pięci lat	pięciu lat
DATIVE	pięci lat	pięciu latom
ACCUSATIVE	pięć lat	pięć lat
INSTRUMENTAL	pięcią lat	pięcioma latami
LOCATIVE	pięci lat	pięciu latach

He further proposes to capture the diachronic change in terms of a grammaticalisation procedure, whereby a higher numeral turns from a content category N to a functional category Q:<sup>7</sup>

- (11) [<sub>DP</sub> D [<sub>NP</sub> pięć<sub>N</sub> [<sub>DP</sub> D [<sub>NP</sub> lat<sub>N</sub> ]]]] Old Polish  
 (12) [<sub>DP</sub> D [<sub>QP</sub> pięć<sub>Q</sub> [<sub>NP</sub> lat<sub>N</sub> ]]] Present Day Polish

This diachronic change leads to both a simplification and complication of the structure of the Q-N relations. The structure is simplified from a bi-nominal frame, with a regular [<sub>DP</sub> D [<sub>NP</sub> N]] content (or a functional projection/lexical projection content) to a single nominal constituent headed by N but insulated by two functional categories [<sub>DP</sub> D [<sub>QP</sub> Q [<sub>NP</sub> N]]].<sup>8</sup> The latter structure becomes more complicated than it used to be. In terms of the feature composition of the N and Q categories, the relevant difference correlates with the presence or absence of the [<sub>person</sub>] feature (cf. 13b) below. Consequently, we propose that the activation of this feature on the higher numeral by T/v is a residue of the diachronic change that took place in the grammar of Polish: activation of this feature finds its source in the past when higher numerals were nouns (cf. 11).<sup>9</sup>

7 Although at some point in our discussion we present nominal structures containing a DP layer, we refrain from taking stance in a discussion on DP/NP status of nominals. What is crucial for us is that nominals in Polish cannot be bare NPs with modifiers in the adjunct positions (contra Bošković 2005). Whether the nominal projection is actually a DP or any other XP (cf. Willim 2000 proposing that Polish nominals are KPs rather than DPs) does not affect the essence of our proposal.

8 For another recent discussion of a historical development of higher numerals in Polish see Miechowicz-Mathiasen (2014).

9 Other accounts of the Genitive of Quantification endorse the dual (adjectival/nominal) character of Slavic numerals (quantifiers) but typically leave it without much discussion (cf. Bošković 2006) or credit its properties to different levels of grammatical representation (D-Structure vs. S-Structure in GB-style theories, cf. Babby 1987 and Franks 1994, 1995).

- (13) a) Adjective [<sub>u</sub>number] [<sub>u</sub>gender] [<sub>u</sub>case]  
 b) Numeral [<sub>u\*</sub>number] [<sub>u\*</sub>gender] ([<sub>i</sub>person]) [<sub>u</sub>case]  
 c) Noun [<sub>i</sub>number] [<sub>i</sub>gender] [<sub>i</sub>person] [<sub>u</sub>case]

- (14) The construction of the Genitive of Quantification is a residue of an earlier, fully nominal stage in the diachronic development of Polish numerals.

Its residual character is clear from its distribution, as it occurs only in a subset of QP environments. It is like the residue of V-2 in English (captured through the feature composition of C), showing only in interrogative and emphatic constructions. Though the [<sub>i</sub>person] feature on a numeral  $\geq 5$  in Present Day Polish is activated by T/v the structure is not bi-nominal as in (11); in other words, the source of the idiosyncrasy of the GoQ construction stems from the fact that the constituent structure of the frame is modern, i.e., (12), but the feature content of numeral from a previous stage is invoked. The in-between pattern in (13b) reflects the Polish and Russian GoQ constructions, where in structural case contexts the high numeral, otherwise adopting the guise of an adjective, puts on the guise of a nominal when matching a finite T/v probe.<sup>10</sup> The change of the guise is due to the activation of an otherwise dormant feature [<sub>i</sub>person]. We also submit that this derivational nominal construct is defective, as it has an impoverished case menu in that both structural cases are lumped into one: accusative.

#### 4. THE NANOSYNTACTIC ACCOUNT OF GENITIVE OF QUANTIFICATION

A sheer number of analyses of the GoQ leads one to believe that this construction serves as a litmus paper for all emerging theories in Slavic linguistics. The number of analyses attempting to explain case properties and distribution within quantified phrases involves a considerable dose of ‘look-ahead’ or a number of countercyclic operations. One of the most typical technical devices involving countercyclic operations is the use of the GB distinction between Deep and Surface Structure operations, where the latter repairs, or fills in the procedural gaps left by the former, with the final grammatical representation in sight. Another popular move to take agreeing numerals ( $<5$ ) to be APs and numerals  $\geq 5$  to be heads, which probably violates **No Tampering** (cf. Chomsky 2000, 2001) when it is to be determined. In Babby (1987) case is assigned by the external head to the maximal projection, i.e., NP, which subsequently spreads to other constituents of the phrase. When Q is present in the structure, in nominative and accusative contexts, it assigns genitive which percolates down to other constituents of the nominal phrase blocking case assignment from the outside. When the external head assigns one of the oblique cases the presence of Q does not hinder case spreading as the inherent case overrides the structural one, hence homogeneous syntax. Some other works, e.g., Franks (1994, 1995) stresses the parametric variation to numeral

10 On the nominal status of higher numerals see also Babby (1987), Greenberg (1978), Corbett (1978a,b) and Caha (2012, 2013) among others.

phrases, i.e., in Russian they are either NPs or QPs, in SC they are NPs and in Polish they present the characteristics of both NPs and QPs. In accusative contexts they are QPs. Przepiórkowski (1999), similarly to Franks, assumes that a higher numeral bears accusative and constitutes the head of the phrase. The noun becomes then its argument, i.e., the subject. This analysis involves considerable look-ahead. In Rutkowski (2002) agreeing numerals, i.e., <5, are viewed as adjectival and thus are introduced in the position of [spec, NP], e.g., (15a), whereas higher numerals are placed in the head position of QP, e.g., (15b):

(15) a)  $[_{DP} [_{D'} D [_{QP} [_{Q'} \text{pięć} [_{NP} [_{N'} \text{osłów}]]]]]]]$   
 ‘five donkeys’

b)  $[_{DP} [_{D'} D [_{QP} [_{Q'} Q [_{NP} \text{dwie} [_{N'} \text{kobiety}]]]]]]]$   
 ‘two women’

Likewise, Bailyn (2004) proposes that the numeral, depending on its value, occupies either the specifier or the head position of QP. The choice of its place in the structure is made at the vP level. The assignment of genitive to the noun is contingent on the position of the numeral, i.e., when it is placed under Q, the case is absorbed and homogenous pattern results, e.g., (16a). With the numeral in [spec, QP], Genitive of Quantification is obtained, e.g., (16b). This relation is formed countercyclically, only after the verb has been merged and the verbal projection has enveloped the NP.

(16) a) The homogeneous pattern:  
 $[_{VP} [_{V'} V [_{QP_{instr}} [_{Q'} \text{pięcioma} [_{NP_{instr}} [_{N'} \text{językami}]]]]]]]$   
 ‘(with) five languages’

b) The heterogeneous pattern:  
 $[_{VP} [_{V'} V [_{Q_{nom/acc}} \text{pięć} [_{Q'} \emptyset [_{NP_{gen}} [_{N'} \text{języków}]]]]]]]$   
 ‘five languages’

Finally, in Bošković (2006, 2013) numerals are lexically specified as either adjectival, hence APs, or quantifiers, i.e., QPs, with agreeing numerals, i.e., <5, occupying the adjunct position of NP, i.e., [FP [F' F [NP AP [N' NP]]]] and Quantifiers, i.e., ≥5 placed in the specifier position of FP, i.e., [FP QP [F' F [NP [N' NP]]]]. Genitive case is assigned by the F head but only when the specifier position of FP is filled by the QP.

Although each of these analyses provide an interesting insight into the nature of QPs, they do not address other intricacies of quantified phrases. Thus, we would like to outline a solution in the spirit of nanosyntax.

The major claim of the nanosyntactic approach is that the sub-word/morpheme level processes are treated in parallel with core syntactic phenomena. In the syntax of nominals, the nanosyntactic model (cf. Starke 2009; Caha 2009, 2010; Taraldsen 2009) provides means to derive various case patterns allowing for movement of the entire NP within the set of case projections (split KP). The analysis of the position of the nominal



head with respect to its satellites (demonstratives, numerals, adjectives) is based on Cinque 2005 which restricts certain types of movement, i.e., rightward and downward movement are forbidden, and determines which chunk of the structure can be moved, i.e., the one with the nominal head. The essential component of the analysis is that particular cases are matched to the functional projections within an articulated Kase Phrase which belongs to the extended projection of the noun (NP). The nominal, in order to acquire a given case, i.e., a proper case suffix, merges in the structure uninflected, and in the course of a derivation moves to the position c-commanding a given case. In our proposal that is the specifier position of a particular case projection. Movement of the nominal is initiated by the probe and its position in the case sequence is determined by language specific constraints. The case sequence and ordering of cases is uniform across languages and is stated in the *Universal Case Contiguity* (Blake 1994; Caha 2009), e.g., (from Caha 2010: 7):<sup>11</sup>

(17) comitative > instrumental > dative > genitive > accusative > nominative [noun]

As example (17) shows, nominative and accusative are placed as the lowest cases in the sequence, which indicates that they are the least marked cases set apart from oblique ones which are usually morphologically more complex (Caha 2009).<sup>12</sup> Also, case syncretisms are predicted to occur on adjacent cases/nodes. Case suffixation follows either an analytic pattern (pied-piping), or a synthetic one. In the former the nominal core (NP) moves successive cyclically to the specifier position of each intermediate case and pied-pipes this functional sequence to its final destination within KP. In the latter the nominal core moves in a single step to its final specifier position (direct movement) and no pied-piping is evident morphologically. In general the nanosyntactic approach to case predicts that nominals in Slavic wear their cases on the sleeve in the sense that the NP moves overtly to a given position within the Kase Projection, which constitutes the external functional projection layer of the nominal constituent.

#### 4.1 Case Projection Sequence and Polish Nominals: Derivation of Genitive of Quantification

In our analysis of Polish nominals we take the noun to be the core element of the phrase, whereas demonstratives, adjectives or numeral quantifiers are located in specifiers and adjoined positions:

11 The case sequence in (17) corresponds to the sequence of NP-external probes that license particular cases, with T licensing nominative, To accusative, D (adnominal) genitive, etc.: ... > D > To > T . A proposal along these lines is put forward in Svenonius (2004).

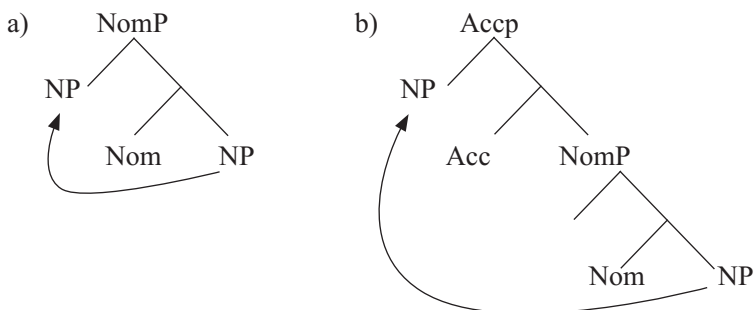
12 The case hierarchy in (17) overlaps to a large degree with the hierarchy proposed in Babby (1987), where the leftmost cases override the rightmost cases on the assumption that lexical properties must be satisfied before the syntactic ones (Principle of Lexical Satisfaction), i.e., Lexical case > GenQ > Nom/Acc.

The sequence of case preference is to be taken representationally, rather than derivationally, so Babby's case overriding is not Pesetsky's (2013) case overwriting.

(18) ... $[_{QP} \text{ Num } [_{Q'} F_Q [_{NP} \text{ Dem } [_{NP} \text{ Adj } [_{NP} N ]]]]]]$

KP, taking active part in the licensing of case, is split into particular case projections and belongs to the extended functional projection of a noun, i.e., there is one articulated KP per a nominal core and its modifiers in Polish. The NP headed by the noun with [+N] feature, moves up to a given position within the case projection sequence, i.e., to the specifier of what we call Nominative Phrase (NomP), Accusative Phrase (AccP) or other case projections where a given case is licensed. The exact motivation for this movement is the need for a successful Spell-Out of a given case suffix, in line with Caha (2009, 2010):

(19)

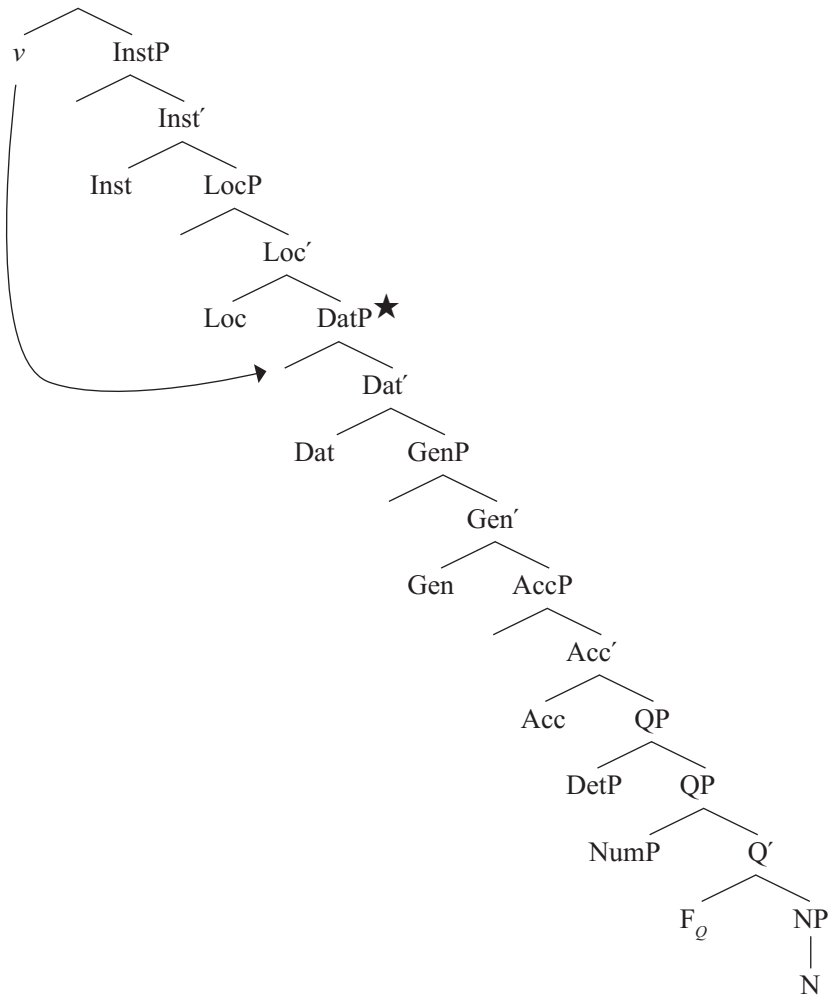


The classic nanosyntactic literature devotes little attention to the relation between the extended projection of the nominal and the split KP and other components of the derivation, becoming liable to the charge of ‘look-ahead’ (cf. Caha 2012, 2013 on numerals). We attempt to incorporate the detailed syntax of case with the syntax of larger components including the nominal (the phrase and the clause) in a manner compatible with the phase theory (e.g., Chomsky 2000, 2001, 2008).<sup>13</sup> Thus in example (20) we present an exemplary derivation with a QP selected by a head, a verb or a preposition, imposing dative on its argument (the homogeneous pattern). We assume that the case paradigm of the higher numeral is defective in that its case projection sequence is truncated at the bottom and NomP is absent, with accusative left as the only structural case.

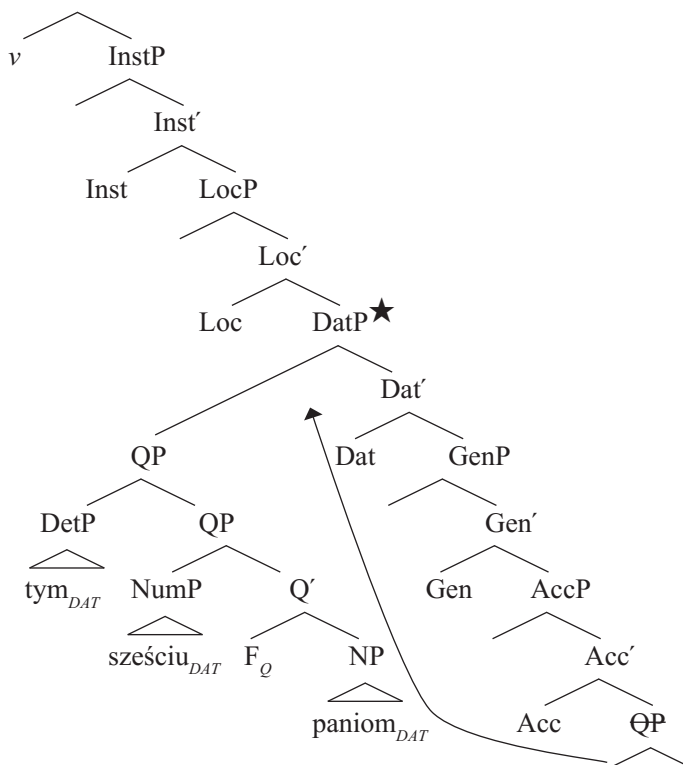
(20) a) *tym*            *sześciu* *paniom*  
           these<sub>DAT</sub>    six<sub>DAT</sub>    ladies<sub>DAT</sub>

13 We propose one KP per a nominal projection, i.e., the head noun and its modifiers, which is a crucial difference between Caha’s (2009, 2010) account, in which projection of every declining element is topped with a separate KP, and ours. This way we can adequately represent the structure of quantified phrases which have changed from a binominal into a single phrase, i.e., a single set of case projections over the QP is a result of the diachronic change discussed in Rutkowski (2007), here examples (11) and (12).

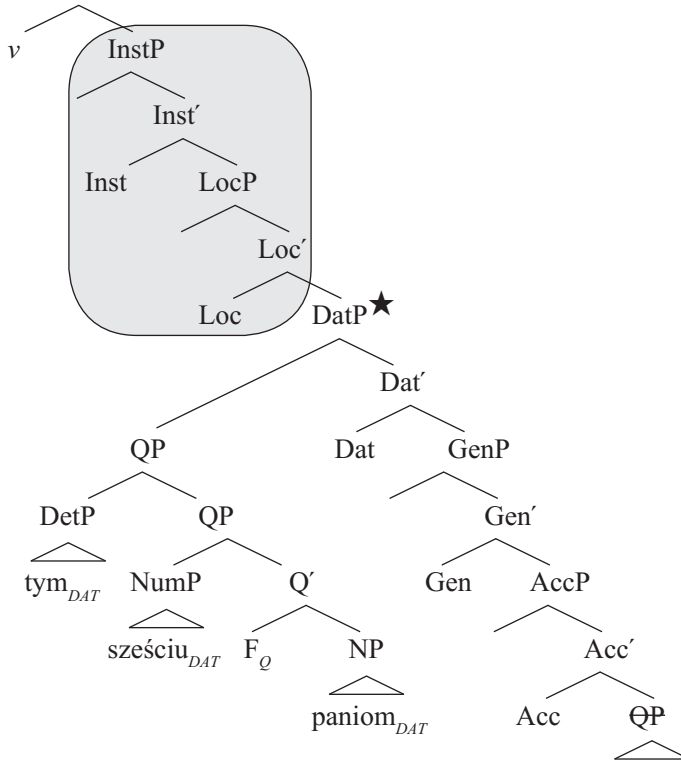
b) activation of DatP



c) movement of QP to DatP



d) deletion of the case projection sequence above DatP



Upon the merger of a (quantified) nominal with a full inventory of cases with the probe,  $v$  or  $P$ , selecting for an argument in a particular case, a relevant case projection becomes activated (20b) and attracts the QP (20c). The movement of QP to [spec, Dat] follows from the postulate of Spell-Out driven movement, whereby the section of the case projection sequence spells out as the dative suffix. As the result, the whole QP phrase moves to the specifier of the Dative Phrase where all the elements in the extended projection of  $N$  (the bearer of the full set of  $\phi$ -features) become marked dative. In the presented derivation, contrary to the major tenet of nanosyntax, the case head affects the entire phrase and the suffix must be appended not only to the final nominal position in the phrase but it must also spread onto the numeral/adjective/demonstrative. At this time we can only propose that the derivational stage in (20) takes place in narrow syntax and the nominal sequence is then marked to be realized with morphological

content on the PF branch.<sup>14</sup> After the part of a derivation in which all the elements of the QP are secured with the right case, other higher and unused projections within split KP become elided, cf. (20d).

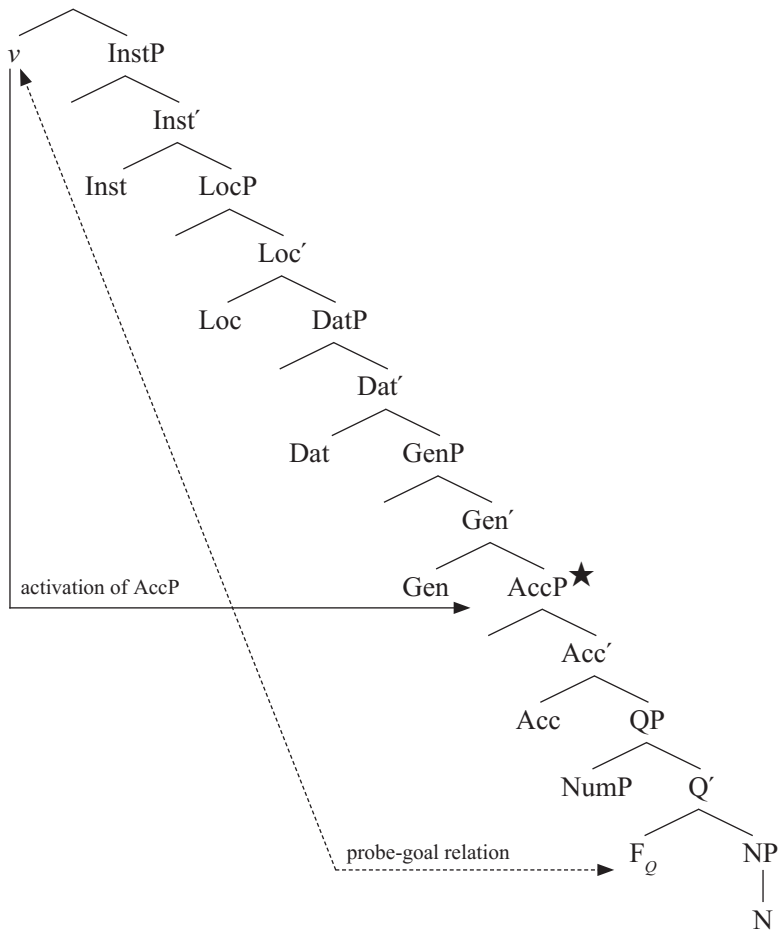
In case of the Genitive of Quantification, the derivation becomes slightly more complex as the quantifier and its nominal complement bear different cases, i.e., accusative and genitive respectively. In order to account for this case mismatch we make two crucial assumptions: (i) within a simple case projection sequence all cases are distinct from one another and their projections are transparent to probing/attraction from higher case heads, without causing minimality/intervention effects; (ii) as the internal morphological composition of certain cases is analytic (i.e., the case suffix of a case higher in the sequence in (20) includes the suffix of a lower case) successive cyclic movement within KP is an option.<sup>15</sup> In such a movement step the Accusative Phrase (accusative marker) can be pied-piped with the NP projection as one constituent. These two elements become instrumental in our account of GoQ. At a certain point in the derivation the v-V complex accesses the case projection sequence with a full set of  $\phi$ -features to value it as accusative and have its own complete  $\phi$ -feature set valued. At the same time Q/F<sub>Q</sub> becomes activated as an adnominal genitive marker and genitive within the split KP becomes activated, see (21b). In the context of nanosyntax we propose the following derivation, where multiple movements within a single KP are crucial. A single case projection sequence over the QP is a result of the diachronic change discussed in Rutkowski (2007).

- (21) a) *Zobaczyłem pięć kobiet.*  
 saw<sub>ISG</sub> five<sub>ACC</sub> women<sub>GEN.PL</sub>  
 ‘I saw five women

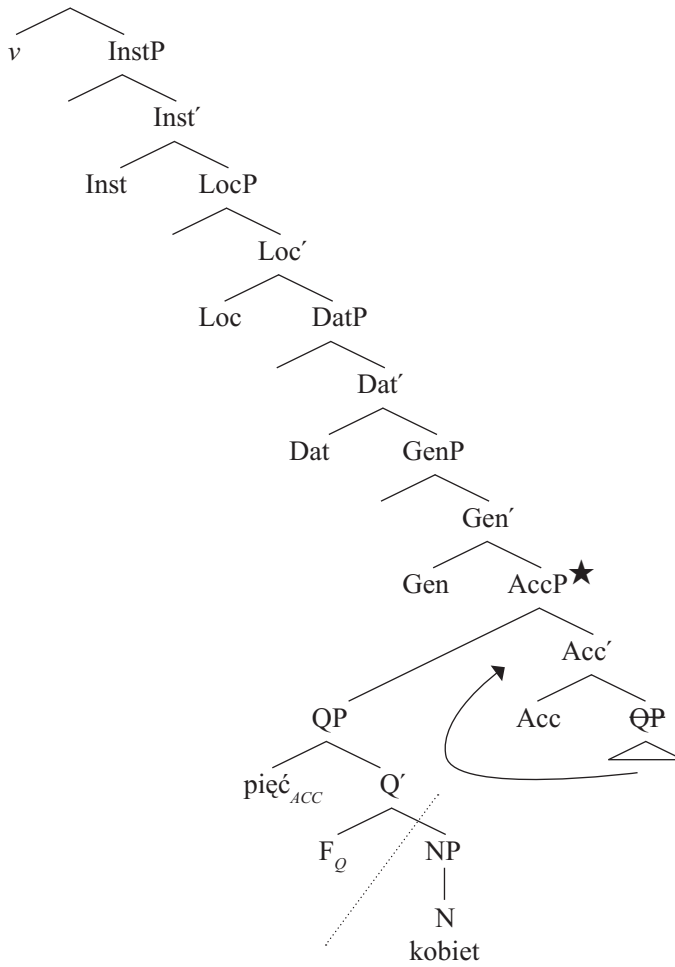
14 By doing so we subscribe to the proposal spelled out in Pesetsky (2013: 99–102) concerning the spread of case within a particular case-marked domain through morphological means. His particular technical solution relies on the use of prototype categories that become sisters to case bearers and has two interesting aspects. First, a prototype  $x^*$  is realized adjacent to the smallest element dominated by the sister of the case licenser. Second, the prototype is not necessarily realized as word-level morphology, but is realized at the lowest structural level that the language and construction permit, which is sometimes phrase-level. In the system developed here, case is appended to the constituent that a given case head attracts and forces it to become its specifier. The lexical realization (for instance spread within this constituent) is determined by the morphology of a given language.

15 Various elements of the same nominal sequence are distinct from each other in the sense of Relativized Minimality (Rizzi 1990) and do not cause intervention effects with respect to one another’s participation in Match, Agree and Move for case. Regarding case licensing on N it may involve pied-piping of its dependents in accordance with Cinque (2005) and the study of permutations in DP/NP internal word orders involving demonstratives, numerals and adjectives, which share a number of properties with case marking viewed as a result of syntactic movement

- b) NumP is activated as a nominal element with a full set of  $\phi$ -features and enters into a probe/goal relation with  $v$ . As a result AccP is activated.

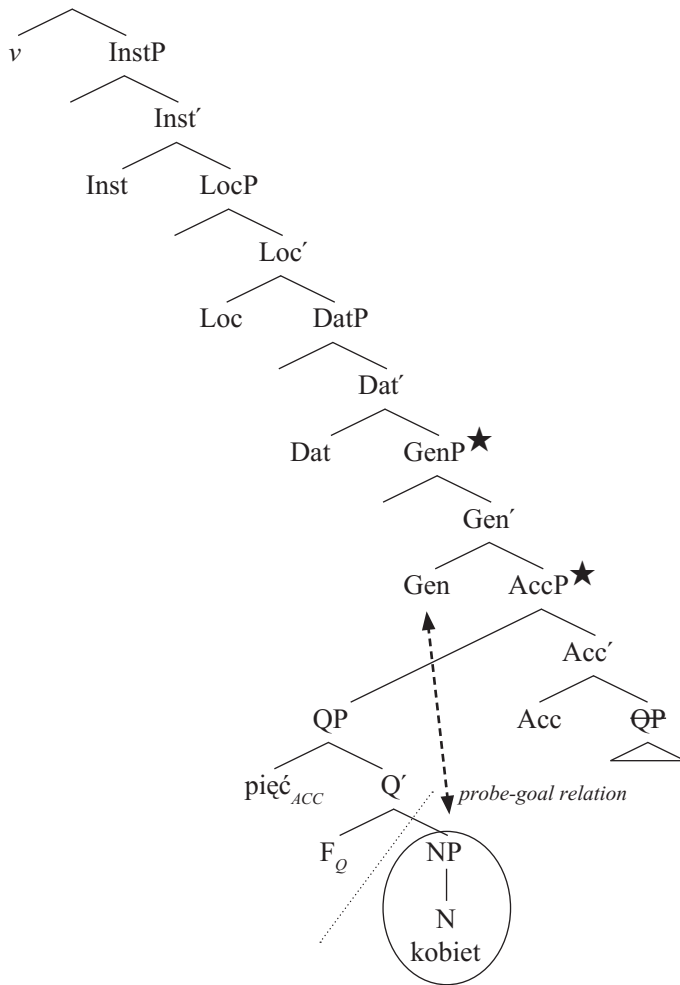


c) QP moves to [spec, Acc].

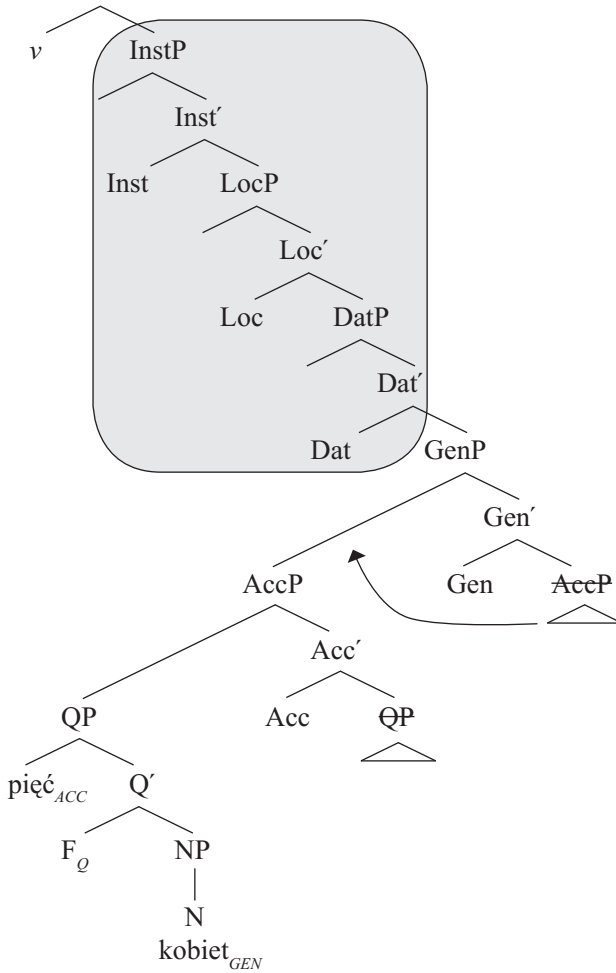




d) NP enters into a probe/goal relation with Gen across QP.



e) NP moves to [spec, Gen] pied-piping [AccP QP] on top. The remaining case projections become deleted.



In the structure presented in (21) the nominal phrase, consisting of a noun, its modifiers and the extended function projection (split KP), is selected as an argument. Then *v* accesses the QP, gets involved in the relation Agree/feature sharing with it and activates AccP in the case projection sequence, cf. (21b). At the same time the higher numeral NumP is accessed by *v*/T<sub>0</sub> and its [<sub>i</sub>person] feature becomes activated, leading to the default setting of all other  $\phi$ -features and turning NumP into an appropriate goal for *v*, as well as allowing it to license a nominal dependent, cf. (21b). As a consequence, the entire phrase (QP including NumP) moves to the position within split KP to appear in the case imposed by the external selector, e.g., *v* makes a QP move to [spec, Acc] (21c).<sup>16</sup> The NP is pied-piped but accusative case is not transmitted to the NP complement which still requires another case.<sup>17</sup> The exceptional nature of this derivation lies in the fact that at this stage of the derivation a single extended functional projection of N (KP) must service two heads bearing independent sets of  $\phi$ -features: the new-born [+ $\phi$ ] NumP and the original [+ $\phi$ ] N. The relation between these two follows an otherwise attested path: a c-commanding nominal [+ $\phi$ ] head forces the other nominal [+ $\phi$ ] head (and its dependents within the same maximal projection) to appear in genitive. This is technically achieved in a Last Resort mode by the Genitive Case Projection, which is activated and accesses the NP across the case marked NumP, cf. (21d). Several reasons come to mind as to why the derivation should allow for this (non-local) relation. First, it takes place within the same extended nominal sequence and no other probe external to the QP is involved. Second, NumP is transparent to the probing from Gen to NP, as its case feature has already been valued. Likewise, the Accusative Phrase is transparent to the attraction of NP by the genitive. Moreover, the Accusative Phrase including NumP is pied-piped in the movement of the NP to its genitive-licensing position, cf. (21e). Third, the derivation in (21) bears the flavor of Richards' 1998 Principle of Minimal

---

16 As pointed out by the reviewer the activation of a person feature and turning a numeral into an element of category  $N(Q_N)$  which is able to value an external case probe (little *v*) as a closer, c-commanding and more minimal goal violates the Inclusiveness Condition. Although we acknowledge this shortcoming of the analysis, it actually follows from the nature of higher numerals which are hybrid, in-between category. Moreover, in order to account for properties of higher numerals either the Inclusiveness Condition is violated or a look-ahead is observed.

17 We must forcefully state that we clearly distinguish between two superficially similar phenomena: case composition and case stacking. The former is ubiquitous in nanosyntax and refers to the morphological composition of case suffixes and is instrumental in establishing the case sequence in example (17) above. It does not presuppose, and must be distinguished from, case stacking understood as a multiple procedure of case marking of one and the same NP set against a number of case licensing heads in the same derivation (cf. Richards 2007; Pesetsky 2013). Case stacking typically involves case overwriting, a procedure of nullifying an earlier case relation [head<sub>1</sub> – NP] by a later relation [head<sub>2</sub> – NP], with or without a morphological trace of the earlier relation showing on the NP. Our account does not presuppose case stacking; on the contrary, we assume that each head bearing a full set of  $\phi$ -features (and its extended projection including dependents, i.e., adjectives and intensifiers) participates in only one case relation per derivation:

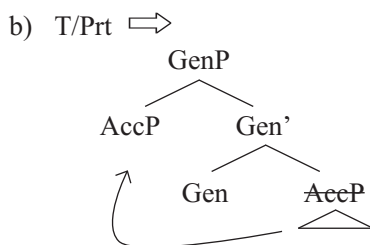
Each head bearing a full set of  $\phi$ -features (and its extended projection including dependents, i.e., adjectives and adjuncts) participates in only one case relation per derivation.

Compliance:<sup>18</sup> within one and the same set of case projections a more local relation is established first (Acc – NumP) before a less local relation is established (Gen – NP). Our account presupposes that there is no case overwriting in Polish (and related languages) and its morphology displays an application of a Genuine Single Suffix Rule: what you see is what you get: Elements placed within QP receive accusative, while elements placed within NP receive genitive.<sup>19,20</sup>

## 5. SOLVING THE AGREEMENT PUZZLE

The nanosyntax inspired account of GoQ leads to a relatively straightforward account of the troublesome agreement patterns mentioned in ex. (2) and (9), repeated below for convenience as (23), with (22) showing the relevant section of the representation:

(22) a) T...Part(icipale)...[<sub>GenP</sub> [<sub>AccP</sub> te pięć kobiet] Gen ]



(23) *Te/tych*      *pięć*    *kobiet*      *było*      *wybrane/wybranych* ...  
 These<sub>ACC/GEN</sub>    five<sub>ACC</sub>    women<sub>GEN.PL</sub>    was<sub>3SG.NEUT</sub>    chosen<sub>3PL.ACC/3PL.GEN</sub>  
 ‘Five women were chosen .....’

The relative configuration of GenP and AccP in (22b) is such that the probe Part is equidistant from both GenP and AccP, on the following assumptions in Pesetsky and Torrego (2001).

18 Principle of Minimal Compliance: For any dependency D that obeys constraint C, any elements that are relevant for determining whether D obeys C can be ignored for the rest of the derivation for purposes of determining whether any other dependency D' obeys C (Richards 1998: 601).

19 We assume that the morphological component on the PF branch of grammar can correctly deal with the marking of both NumP in spec,QP with accusative and N (NP) with genitive, on the assumption that both head/spec relation (accusative) and linear adjacency (genitive) are legitimate relations for morphology to operate on.

20 One of the consequences of our account is that the default adnominal case must be higher within the KP sequence than the initial structural case absorbed by the nominalized numeral:

The default adnominal case postulate: The default adnominal case projection is placed higher in the case hierarchy than structural cases.

- (24) a) Attract Closest: If a head K attracts feature F on X, no constituent that bears F is closer to K than X.  
 b) Closeness: Y is closer to K than X if K c-commands Y and Y c-commands X.

In their discussion of the that-trace effect, they argue that in the configuration below both the specifier of T and the projection of TP can delete the (same) feature uT on C:

- (25) We know [<sub>CP</sub> [<sub>uT</sub>] [<sub>TP</sub> [the student<sub>uT</sub>] T{that<sub>IT</sub>} [<sub>VP</sub> bought the book]]]

The notion of closeness is a crucial component of the relations of not only Attract (and Move) but also Match and Agree in the phase-based theory (Chomsky 1995, 2000, 2001, 2008). For instance, the probe is supposed to match and agree with the closest potential goal. In (20) GenP dominates AccP, so it does not c-command it, and while AccP c-commands Gen, the same label as the label on Gen is present on GenP. In the context of our discussion of GoQ in the nanosyntax inspired framework, the structure of relevant QPs in Polish and Russian looks as follows:

- (26) a) T [<sub>GenP</sub> [<sub>NomP</sub> ...] ... Gen] Rus.  
 b) T [<sub>GenP</sub> [<sub>AccP</sub> ...] ... Gen] Pol/Rus.

Russian allows the variants (26a-b), while Polish allows only for (26b). In Polish whenever the probe T cannot find a nominal goal that is marked for Nom its  $\phi$ -features default to 3<sub>SG,NEUT</sub>. Though the probe T has a choice of two close(r) goals neither can value its  $\phi$ -features and T defaults. In Russian two subject/verb concord possibilities are available but they are not fully equivalent. We assume that the Agree operation in which the  $\phi$ -features of T are fully valued is more economical (and generally preferable) to the option in which they default. Say that defaulting involves an extra derivational step and incurs extra burden on the derivation (cf. Preminger 2009). Thus whenever the QP has the structure in (26a), T shows full agreement in Russian. In the cases of default agreement the Russian QP shares its structure with its Polish equivalent. As for the puzzling agreement with the head of the Participle Phrase or predicative adjective in Polish, the relevant configuration is the same. The probe needs to agree for case with either goal that bears this feature; as it happens two candidates are locked in this configuration:

- (27) Part [<sub>GenP</sub> [<sub>AccP</sub> te pięć dziewczyn...] Gen ]

Unlike T, Part has an incomplete  $\phi$ -probe, which misses the [<sub>person</sub>] feature and functions as a passive recipient of the features provided by its nominal goal and probe T. A default T makes no claims on the features of Part, whereas both GenP and AccP are close to Part on the strength of (22), providing it with a free option.<sup>21</sup> Therefore Part

21 The variation within the case marking of a demonstrative (i.e., accusative or genitive, qualifying the numeral only or the entire QP, see fn.6 for relevant examples) as well as its scope does not correlate with the case on the participle. What matters here is the timing of adjunction of a de-



(pied-piping the QP above it) is raised to [spec, Gen]. The double satisfaction of the case requirements produces a structure of a distinct potential for case feature checking of the participle and predicative adjective.

## References

- BABBY, Leonard (1987) "Case, prequantifiers, and discontinuous agreement in Russian." *Natural Language and Linguistic Theory* 5, 91–138.
- BAILYN, John F. (2004) "The case of Q." In: Olga Arnaudova (ed.), 1–36.
- BLAKE, Barry (1994) *Case*. Cambridge University Press: Cambridge.
- BOŠKOVIĆ, Željko (2005) "On the locality of left branch extraction and the structure of NP." *Studia Linguistica* 59, 1–45
- BOŠKOVIĆ, Željko (2006) *Case and agreement with genitive of quantification in Russian*. In: Cedric Boeckx (ed.), 99–120.
- BOŠKOVIĆ, Željko (2013) *More on the edge of the edge*. In: Cassandra Chapman/Olena Kit/Ivona Kucerova (eds), 44–66.
- CAHA, Pavel (2009) *Nanosyntax of case*. PhD Dissertation. University of Tromsø.
- CAHA, Pavel (2010) *The parameters of case marking and spell out driven movement*. <http://ling.auf.net/lingbuzz/001026>.
- CAHA, Pavel (2012) *Czech numerals*. Paper presented at Cartography – Where do we go from here?, Geneva, June 7–9, 2012.
- CAHA, Pavel (2013) *Czech numerals and no bundling*. <http://ling.auf.net/lingbuzz/001723>.
- CINQUE, Guglielmo (2005) "Deriving Greenberg's Universal 20 and its exceptions." *Linguistic Inquiry* 36, 315–332.
- CHOMSKY, Noam (1995) *The Minimalist Program*. Cambridge, MA: the MIT Press.
- CHOMSKY, Noam (2000) *Minimalist inquiries*. In: Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta (eds) *Foundational issues in linguistic theory*. Cambridge, MA: MIT Press, 89–156.
- CHOMSKY, Noam (2001) "Derivation by phase." In: Michael Kenstowicz (ed.), *Ken Hale. A Life in Language*. Cambridge, Mass.: The MIT Press, 1–52.
- CHOMSKY, Noam (2008) *On Phases*. In: Robert Freidin/Carlos P. Otero/Maria Luisa Zubizarreta (eds), 133–166.
- CORBETT, Greville G. (1978a) "Numerous squishes and squishy numerals in Slavonic." *International Review of Slavic Linguistics* 3, 43–73.
- CORBETT, Greville G. (1978b) "Universals in the syntax of cardinal numerals." *Lingua* 46, 355–368.
- GREENBERG, Joseph. H. (1978) *Generalizations about numeral systems*. In: Joseph H. Greenberg (ed.), *Universals of Human Language*. Vol 3. Stanford: Stanford University Press, 249–295.
- FRANKS, Steven (1994) "Parametric properties of numeral phrases in Slavic." *Natural Language and Linguistic Theory* 12, 597–674.
- FRANKS, Steven (1995) *Parameters of Slavic morphosyntax*. New York: Oxford University Press.

- FRANKS, Steven (2002) "A Jakobsonian feature based analysis of the Slavic Numeric Quantifier Genitive." *Journal of Slavic Linguistics* 10, 141–181.
- MIECHOWICZ-MATHIASSEN, Katarzyna (2014) *Numeralization of numeral nouns in Polish*. In: Ludmila Veselovská/Markéta Janebová (eds), 48–68.
- PERELTSVAIG, Asya (2006) "Small nominals." *Natural Language and Linguistic Theory* 24, 433–500.
- PESETSKY, David (1982) *Paths and Categories*. PhD Dissertation. *Massachusetts Institute of Technology*.
- PESETSKY, David (2013) *Russian case morphology and the syntactic categories*. Cambridge: The MIT Press.
- PESETSKY, David/Esther TORREGO (2001) *T-to-C movement: Causes and consequences*. In: Michael Kenstowicz (ed.), 355–426.
- PREMINGER, Omer (2009) *Failure to Agree is not a failure:  $\phi$ -Agreement with post-verbal subjects in Hebrew*. Manuscript. MIT/ Harvard.
- PRZEPIÓRKOWSKI, Adam. (1999) *Case assignment and the complement-adjunct dichotomy: A non-configurational constraint-based approach*. PhD Dissertation. Universität Tübingen.
- PRZEPIÓRKOWSKI, Adam (2004) "O wartości przypadku podmiotów liczebnikowych." *Bulletin De La Société Polonaise De Linguistique, fasc. LX*, 133–143.
- PRZEPIÓRKOWSKI Adam/Agnieszka PATEJUK (2012) *The puzzle of case agreement between numeral phrases and predicative adjectives in Polish*. In: Miriam Butt/Tracy H. King (eds), 490–502.
- RIZZI, Luigi (1990) *Relativized Minimality*. Cambridge: The MIT Press.
- RICHARDS (1998) "The Principle of Minimal Compliance." *Linguistic Inquiry* 29/4, 599–629.
- RICHARDS, Norvin (2007) "Lardil 'Case Stacking' and the structural/inherent case distinction". Unpublished manuscript. *Massachusetts Institute of Technology*.
- RUTKOWSKI, Paweł (2002) "The syntax of quantifier phrases and the inherent vs. structural case distinction." *Linguistic Research* 7/1, 43–74.
- RUTKOWSKI, Paweł (2007) *Hipoteza frazy przedimkowej jako narzędzie opisu składniowego polskich grup imiennych*. PhD Dissertation. Uniwersytet Warszawski.
- STARKE, Michal (2009) "Nanosyntax: A short primer to a new approach to language." *Nordlyd* 36/1, 1–6.
- SVENONIUS, Peter (2004) *On the edge*. In: David Adger/Cécile de Cat/George Tsoulas (eds), 261–287.
- TARALDSEN, Tarald (2009) "Lexicalizing number and gender in Lunigiana." *Nordlyd* 36, 113–127.
- WILLIM, Ewa (2000) "On the grammar of Polish nominals." In: Roger Martin/David Michaels/Juan Uriagereka (eds), 319–346.
- WITKOŚ, Jacek/Dominika DZIUBAŁA-SZREJBROWSKA (2015) "A note on the Genitive of Quantification in Polish and derivational phrases." *Poznań Studies in Contemporary Linguistics* 51(3), 433–462.



WITKOŚ, Jacek/Dominika DZIUBAŁA-SZREJBROWSKA (2016) "Numeral phrases as subjects and agreement with participles and predicative adjectives." *Journal of Slavic Linguistics* 24(1), 225–260.

#### Summary

#### SOME ASPECTS OF AGREEMENT WITH NUMERAL PHRASES IN POLISH

The aim of this article is to briefly analyze the agreement patterns in Polish constructions with quantified subjects and adjectival predicates/participles, and propose an account built on the nanosyntactic ideas regarding the nature of case, i.e., split Kase Phrase (Caha 2009, 2010). In the analysis we address the troublesome issues regarding the Genitive of Quantification, i.e., the source of Genitive on the nominal complement in structural contexts, and the optionality in agreement in case between the adjectival predicate/participle and the numeral ( $\geq 5$ ), or the noun of the quantified subject. The essential part of our proposal is based on the articulated Kase Phrase in the functional sequence of the extended nominal projection and its role in the syntactic derivation of case in the spirit of nanosyntactic approach.

**Keywords:** adjectival and participial agreement, Genitive of Quantification, split KP, quantified subjects

#### Povzetek

#### NEKATERI VIDIKI UJEMANJA S ŠTEVNIŠKIMI ZVEZAMI V POLJŠČINI

Namen članka je kratko analizirati vzorce ujemanja v zgradbah s kvantificiranim osebkom in pridevniškim predikatom/deležnikom v poljščini ter predlagati razlago, ki temelji na nanoskladenjskem pristopu k naravi sklona, tj. na podlagi deljene sklonske zveze (Caha 2009, 2010). V analizi se ukvarjamo s problematičnimi vidiki kvantifikacijskega roditelja, tj. z izvorom roditelja na samostalniškem dopolnilu v strukturalnih kontekstih ter s poljubnostjo ujemanja v sklonu med pridevniškim predikatom/deležnikom in števnikom ( $\geq 5$ ) ali samostalnikom kvantificiranega osebk. Bistveni del naše teorije temelji na artikulirani sklonski zvezi znotraj funkcijskega niza razširjene samostalniške projekcije in njeni vlogi pri skladijski derivaciji sklona v duhu nanoskladenjskega pristopa.

**Ključne besede:** pridevniško in deležniško ujemanje, kvantifikacijski roditelj, deljena sklonska zveza, kvantificirani osebek



## CONDITIONAL ANTECEDENTS CONTAINING THE GERMAN DISCOURSE PARTICLE *DENN*: A CORPUS STUDY\*\*\*

### 1. INTRODUCTION

The German expression *denn* has various functions, for example, as a sentence connector, as a comparative particle, and notably as a discourse particle. The literature on discourse particle *denn* nearly exclusively discusses its use in questions (see Thurmair 1989, 1991; Bayer 2012, i. a.). In this paper, we take a look at a second, understudied use of discourse particle *denn*, as in (1), which occurs in the antecedent of a conditional (henceforth: conditional *denn*; see Brauße 1994; Kwon 2005; Coniglio 2011; Häussler 2015).

- (1) [Context: A asks B which activities are planned for the next day.]  
B: *Wir gehen schwimmen, wenn es denn warm genug ist.*  
we go swimming if it DENN warm enough is  
B: 'We'll go swimming, if it is DENN warm enough.'

By using conditional *denn*, B signals that she is uncertain, even skeptic, that it will be warm enough to go swimming the next day. Hence, *denn* intuitively strengthens the pragmatic inference connected to the antecedent<sup>1</sup> that the speaker does not believe that the proposition denoted by the antecedent (i.e., the “antecedent proposition”) holds.

The aim of this paper is to present two corpus studies that shed light on one of the conditions of use of conditional *denn* by exploring the behavior of antecedents containing *denn* (henceforth: AWD), and to discuss a classification of the corpus data based on observations from the studies.

\* sarah.zobel@ds.uni-tuebingen.de

\*\* eva.csipak@uni-konstanz.de

\*\*\* We thank Edith Scheffell for bringing the use of the particle *denn* in conditionals to our attention. We also thank Andrea Beltrama, María Biezma, Ryan Bochnak, Kai von Fintel, Viola Schmitt, Frank Sode, Yvonne Viesel, Thomas Weskott, and the audiences at CSSP 2015, SinFonJA 8, Questions in Discourse 7, and María Biezma's course “Pragmatics 3” (Konstanz, Summer 2016) for helpful discussion. We also thank an anonymous reviewer for comments on the manuscript. All mistakes are our own.

1 For the sake of brevity, whenever we use the term “antecedent”, we mean conditional antecedents.

The paper is structured as follows. We briefly present the existing literature on conditional *denn* and introduce our own account of conditional *denn*'s conditions of use and semantic contribution in Section 2. In Section 3, we present the results of our corpus studies on the behavior of AWDs and discuss their implications. The classification of the corpus data is presented in Section 4. Section 5 concludes.

## 2. A SEMANTIC/PRAGMATIC ANALYSIS OF CONDITIONAL *DENN*

### 2.1 The Function of Discourse Particles and Previous Analyses of Conditional *denn*

The function of discourse particles is to fit the current utterance to the ongoing discourse (Zimmermann 2011). That is, they convey speaker attitudes or additional information on the content of that utterance. In this sense, they are “discourse navigating devices” (see McCready 2006; Eckardt 2013; Rojas-Esponda 2014) that are utilized by the speaker to make explicit certain parts of the discourse, for instance, the make-up of the current “common ground” (= the beliefs the speaker and her addressee share as a result of their conversation, Stalnaker 1973). It is commonly assumed that formally, discourse particles contribute “not-at-issue content”, that is, content that is not part of the truth conditions of the sentence that contains the particles (see Simons et al. 2011; Potts 2015).

As we have stated in the introduction, conditional *denn* is understudied compared to *denn* in questions.<sup>2</sup> In the literature, conditional *denn* is discussed in Brauße 1994, Kwon 2005, Coniglio 2011, and Häussler 2015. Differing in the details, these authors agree that conditional *denn* seems to signal the speaker's uncertainty or doubt about the truth of the antecedent proposition. While intuitively appealing, this leaves open the question as to how the contribution of *denn* differs from and interacts with the inference connected to the antecedent that the speaker is not committed to the truth of the antecedent in the actual world. Our proposal, which we give below, is completely explicit regarding this point.

### 2.2 A New Analysis of Conditional *denn*

Below, we summarize our own analysis of the conditions of use and the semantic contribution of conditional *denn*. For reasons of space, we cannot fully motivate and discuss our analysis; for details, we refer the interested reader to Csipak and Zobel (to appear).

The three parts of our proposal in (2)-(4) are based on our native speaker intuitions regarding constructed examples and a sample of naturally occurring data taken from the ZEIT corpus from the online platform “Digitales Wörterbuch der Deutschen Sprache” (“Digital Dictionary of the German Language”, DWDS).

---

2 Discourse particle *denn* in questions has been discussed quite extensively while conditional *denn* has been mostly overlooked. This is not surprising given the relative frequency of the two uses: compared to *denn* in questions, conditional *denn* is rare; the latter use makes up only about 3-5% of all particle uses. This estimate is based on a random sample of 200 tokens of *denn* exported from the corpus of Spoken German (“Gesprochene Sprache”, ~2.5 million tokens) that is part of the DWDS platform (export: Jan. 30, 2016). For reasons of space, the details of this study cannot be presented.

The first condition of use for conditional *denn* given in (2) captures that *denn* can only be used if the speaker is uncommitted to the truth of the proposition  $p$  that *denn* comments on.

(2) Condition 1

The speaker does not believe that  $p$  is true in the actual world  $w_0$ , that is, he is uncommitted to the truth of  $p$  in  $w_0$ .

Antecedents of hypothetical conditionals fulfill this condition while antecedents of temporal and factual conditionals do not (see Fabricius-Hansen and Sæbø 1983; von Fintel 2011). Above, we observed that conditional *denn* makes explicit the speaker's uncertainty regarding the antecedent proposition  $p$  for the antecedent in which it occurs. When is a speaker prompted to express uncertainty or doubt with respect to a proposition  $p$ ? For example, when there is evidence in the discourse context that someone acts as if  $p$  were true, but the truth of  $p$  is not supported by the speaker's knowledge about the actual world. Condition 2 restricts the use of *denn* to a subtype of this kind of context: the proposition  $p$  on which *denn* comments must have been tacitly proposed.<sup>3</sup>

(3) Condition 2

The proposition  $p$  is tacitly proposed or can reasonably be inferred to be tacitly proposed by a participant  $\alpha$ , where  $p$  is a necessary precondition for the validity of the content of a previous utterance by  $\alpha$  (or a part of that utterance).<sup>4</sup>

Any non-explicitly conveyed content qualifies as a tacit proposal. The term “necessary precondition” is not meant in a logical sense; it includes presuppositions, as well as premises of defeasible inferences based on world knowledge. In sum, the concept of “tacitly proposed necessary precondition” is a generalization of the notion of presupposed new information. This is information that the speaker asks the addressee to accommodate before regular discourse updates can be performed (see von Fintel 2008).

Lastly, we assume that the not-at-issue content contributed by the particle interacts with the pragmatic inference accompanying the use of a hypothetical conditional: the speaker is uncommitted to the truth of the antecedent proposition  $p$  in  $w_0$  (as required by Condition 1). By adding *denn*, the speaker signals that she is not only uncommitted to the truth of  $p$ , but in fact judges  $p$  as so improbable that she would not be willing to assert it.

(4) Not-at-issue content contributed by *denn*.

$[[denn]]^c(p): \text{prob}(c_s, w_0, p) < T$ ,  
where  $T$  is the threshold for assertability

In prose: The probability assigned by the speaker  $c_s$  to whether  $p$  is true in  $w_0$  is less than a pragmatic assertability threshold  $T$ .

3 Condition 2 also excludes *denn* from occurring in factual conditionals.

4 Note that  $\alpha$  can also be the speaker.

Example (5) illustrates the entire proposal.

(5) [Context: Speaker A discusses his first visit to his fiancé's brother with a friend.]

A: *Sein Auto habe ich nicht gesehen, wenn er denn eines hat.*

his car have I not seen if he DENN one has

A: 'I didn't see his car, if he DENN has one.'

Condition 1 is met because *denn* occurs in the antecedent of a hypothetical conditional. Condition 2 is also met. The definite description *sein Auto* (Engl. 'his car') in the consequent presupposes that A's fiancé's brother has a car (= *p*). Since it is A's first visit to the brother's house, the presupposed content *p* is plausibly not part of the common ground. Hence, by using the definite description, A makes the tacit proposal to update the common ground with '*He has a car*'. Lastly, the use of conditional *denn* signals that speaker A is uncertain/skeptic that his fiancé's brother has a car, and that he is, hence, unwilling to assert or to presuppose *p*.

While Condition 1 captures a precise property of conditional *denn*, the exact nature of Condition 2 raises more questions. To gain a clearer empirical picture, we conducted two corpus studies to investigate whether there is evidence for the presence of a preceding tacit proposal. The results are presented in Section 3.

### 3. CORPUS STUDIES ON CONDITION 2

#### 3.1 Study 1

##### 3.1.1 Operationalization

The obvious problem with finding an adequate operationalization for pragmatic concepts like "presence of a tacit proposal" is how to translate them into categories that can be reliably and potentially automatically checked for in a sample of corpus data. The operationalization that we use in our first study is "position of the antecedent relative to its consequent". While the items cannot be tagged automatically for this property, reliable annotation criteria can be given easily.

The motivation for this choice is the following observation: If a speaker uses AWDs to express uncertainty with respect to a tacit proposal made by her interlocutor, our intuition is that she preferably uses a bare antecedent. In contrast, if the speaker uses an AWD to self-qualify her own statements, she preferably uses a full conditional where the tacit proposal occurs in the consequent.

(6) Exploratory hypothesis 1

Antecedents containing *denn* in full conditionals occur more frequently following an overt consequent than preceding it.<sup>5</sup>

---

5 The conditions of use for discourse particles are not strict rules (cf. Zimmermann 2011: 2027). Hence, we can only expect to find the predicted patterns regarding Condition 2 in the majority of cases.

### 3.1.2 Choice of Corpus and Query

We chose the ZEIT corpus available at the DWDS platform. ZEIT is a corpus of journalistic texts (~225.8 mio. tokens, 2015) taken from the German weekly newspaper DIE ZEIT, which among traditional articles also contains transcribed interviews (i.e., texts in written language that is close to spoken language), which arguably increases the possibility of encountering discourse particles. To specifically target AWDs, we used the query in (7).

(7) (“wenn #7 denn”) || (“@falls #7 denn”) && !(“\, denn”) && !(“\; denn”) && !(“\: denn”)

The query includes *denn* following the conditional subjunctors *wenn* or *falls*, and excludes the occurrences of *denn* introducing an independent clause.<sup>6</sup> The query (May 3, 2015) yielded 4,411 results which were all exported. From this sub-corpus, we randomly picked a sample of 300 items, which were filtered manually for false hits and problematic items.<sup>7</sup> The final number of corpus items that were annotated was 260.

### 3.1.3 Details Regarding the Annotation

To annotate the relative position of the antecedent, we used the two field categories “Vorfeld” and “Nachfeld” of the classical Topological Field Model (see Höhle 1986): the “Vorfeld” contains the linguistic material before the finite verb in German main clauses; the “Nachfeld” contains the material after the non-finite verbal material in main clauses and all verbal material in embedded clauses. In addition to these categories, we used the levels “bare” for bare antecedents and “parenthetical” for antecedents that are inserted parenthetically after the Vorfeld constituent or as part of the main body of the sentence (= the “Mittelfeld”).<sup>8</sup> The latter level was not planned initially, but proved to be necessary during the annotation process.<sup>9</sup>

(8) POSITION OF ANTECEDENT (4 levels)

bare (9a), Vorfeld (9b), parenthetical (9c), Nachfeld (9d)

(9) a) *Wenn sie das **denn** angegeben und öffentlich zugänglich gemacht haben.*

if they that DENN specified and publicly accessible made have  
‘If they DENN specified it, and made it publicly accessible.’\*

b) *Wenn’s **denn** so polar zugehen sollte, dann ziehe ich die USA vor.*

if-it DENN so polar be-like-this should then prefer I the USA PRT  
‘If we are DENN supposed to have a strict opinion, then I prefer the US.’\*

6 The conjunction *denn* expresses a causal link between the two conjuncts.

7 Items were classified as problematic if they were fragments, or if it was impossible for us to make sense of the content.

8 In the Nachfeld, AWDs that are semantically embedded and parenthetical AWDs were not distinguished. For the purpose of this study, the distinction is not crucial.

9 Examples that were taken from the corpus sample (possibly with slight editing) are marked by \* after the translation.

- c) *Aber dazu war Neuber, wenn er **denn** gewollt hätte, gar nicht in der Lage.*  
 but to-this was Neuber if he DENN wanted had at-all not in the position  
 ‘But Neuber was not in the position to do this if he would DENN have wanted to.’\*
- d) *Ein Handy, mit dem man alles kann, wenn man es **denn** könnte.*  
 a cell-phone with which one everything can if one it DENN could  
 ‘A cell phone with which you can do everything if you are DENN able to.’\*

We annotated the 260 items independently using the criteria above; we achieved a percentage of agreement of 91.9% ( $\kappa = 0.88$ ). In a second step, we discussed the items for which we had disagreed in our separate annotations and decided on a final annotation for the quantitative evaluation.

### 3.1.4 Results and Discussion

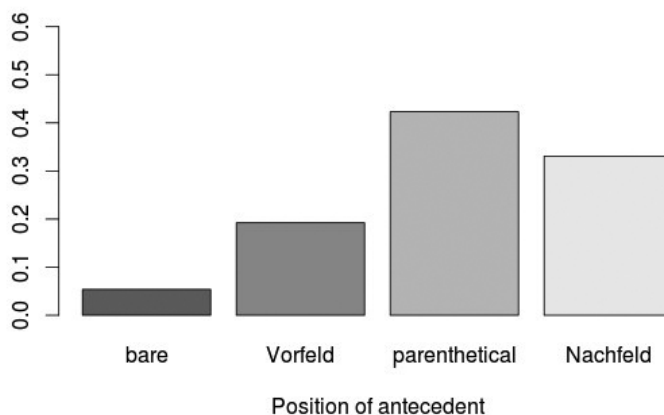
The frequencies of the four levels of POSITION OF ANTECEDENT found in the sample are given in Table 1; the relative frequencies are illustrated in Figure 1.

Table 1: Frequencies POSITION OF ANTECEDENT – AWDs

	BARE	VORFELD	PARENTHETICAL	NACHFELD
FREQUENCIES	14 (5.4%)	50 (19.2%)	110 (42.3%)	86 (33.1%)

The frequencies of the four levels differ significantly ( $\chi^2 = 81.42$ ,  $df = 3$ ,  $p < 0.05$ ); the direct comparison of the levels “Vorfeld” and “Nachfeld” also reveals a statistically significant difference ( $\chi^2 = 9.53$ ,  $df = 1$ ,  $p < 0.05$ ). Hence, the data support our hypothesis.

Figure 1: Relative frequencies POSITION OF ANTECEDENT – with *denn*



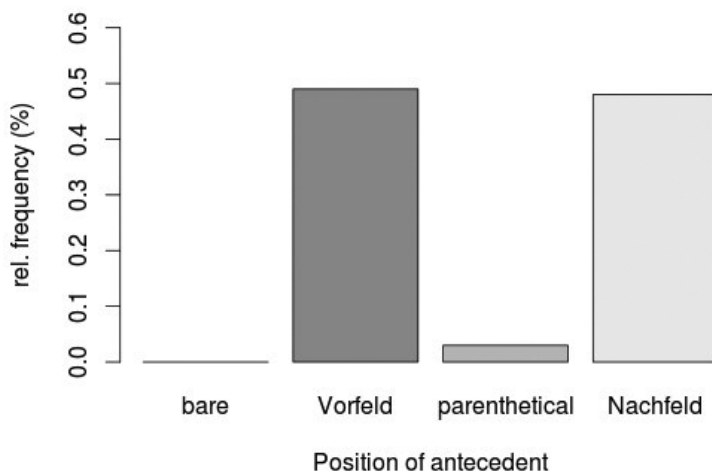
The high number of parenthetical AWDs was not anticipated. Parenthetical AWDs come with the additional complication that the relative position with respect to the tacit proposal given in the consequent has to be determined on a case-by-case basis. This means that the placement of an AWD relative to the consequent cannot be used as a heuristic for the relative position of the AWD and the tacit proposal. Since parenthetical AWDs form the largest group, our original operationalization is less reliable than we initially assumed. A different method was chosen for Study 2.

Given the surprising amount of parenthetical AWDs, we investigated whether antecedents of hypothetical conditionals in general behave like AWDs regarding their syntactic positions. We annotated 100 items of hypothetical conditionals (ZEIT corpus, DWDS) for the same factor and levels (agreement: 98%,  $\kappa = 0.96$ ). The result is given in Table 2 and Figure 2.<sup>10</sup>

Table 2: Frequencies of POSITION OF ANTECEDENT – hyp. conditionals

	BARE	VORFELD	PARENTHETICAL	NACHFELD
FREQUENCIES	0 (0%)	48 (48%)	3 (3%)	49 (49%)

Figure 2: Relative frequencies of POSITION OF ANTECEDENT – hyp. conditionals



A direct comparison of Figure 1 and Figure 2 clearly shows that AWDs behave quite differently from antecedents of regular hypothetical conditionals. This suggests that there is indeed a factor that influences placement of AWDs. We argue that this factor is the presence of a preceding tacit proposal to which conditional *denn* is sensitive.

<sup>10</sup> Our study reproduces the results in Volodina (2014: 756ff).



Taking a closer look at the placement of a handful of parenthetical cases, we also observed that AWDs are preferably placed as close to the perceived source of the proposal as (syntactically) possible. This is further supported by the descriptive results of Study 2.

## 3.2 Study 2

### 3.2.1 Items and Operationalization

Since the first syntactic operationalization was shown to lack the necessary precision, we chose to annotate the same sample items as in Study 1 (based on our informed native speaker intuitions) regarding the relative positions of the AWD and the material that we identify as the source of the tacit proposal:<sup>11</sup>

- (10) Exploratory hypothesis 2  
Antecedents containing *denn* more often follow the source of the tacit proposal than precede it.

### 3.2.2 Details Regarding Annotation

We annotated two factors, RELATIVE POSITION OF PROPOSAL and PROPOSAL IN SENTENCE. The levels of the first factor answer the question “Where does the tacit proposal originate in the discourse relative to the position of the AWD?” The final set of levels for this factor was determined during the annotation process.

- (11) RELATIVE POSITION OF PROPOSAL (5 levels)  
before, around, after, inside, indeterminable

The levels “before” (12) and “after” (13) encode the following: in case we could identify the lexical material to which we could ascribe the source of the tacit proposal, this material (inside the preceding context or the item itself) was entirely before or after the AWD.<sup>12</sup>

- (12)  
*Ein Handy, mit dem man alles kann, wenn man es **denn** könnte.*  
a cell-phone with which one everything can if one it DENN could  
'A cell phone with which you can do anything if you are DENN able to.'\*

- (13)  
*Wenn sie **denn** nicht vermeidbar ist, wie sollte eine Kündigung kommuniziert werden?*  
if she DENN not avoidable is how should a termination communicated become  
'If it is DENN not avoidable, how should a termination be communicated?''\*

---

11 We are aware of the problems connected with annotating non-objective properties of data. We believe, however, that new insights can be gained this way, nevertheless, given the exploratory nature of this study and adequate caution.

12 The lexical source material of the tacit proposal which we identified is underlined.

The level “around” (14a) was assigned to items for which the lexical source material was partly before and after the antecedent; the level “inside” (14b) was given to items for which the antecedent explicitly repeated (part of) the lexical source material.

(14)

- a) *Wohin* aber *soll* *sich die Partei*, wenn sie es *denn* überhaupt will, *wenden*?  
 where but shall SELF the party if she it DENN at-all wants turn-to  
 ‘But which new orientation should the party choose if it DENN wants to’\*  
 b) *Und wenn es denn ein Vorwurf sein sollte*, *er sei ja ein deutscher Patriot*...  
 and if it DENN a reproach be should he is PRT a German patriot  
 ‘And if it should DENN be a reproach that he is a German patriot...’\*

Lastly, the level “indeterminable” was used for items for which no plausible lexical source material could be determined inside the preceding context (one sentence) or the item itself.

The factor REL. POSITION OF PROPOSAL does not encode information about whether the source of the proposal can be found in the item or the preceding context. Hence, we annotated the second factor PROPOSAL IN SENTENCE. Its two levels are the answers to the question “Is there evidence that the tacit proposal originates inside the conditional?” The level “yes” is illustrated, for instance, in (12), “no” is illustrated in (16).

(15) PROPOSAL IN SENTENCE (2 levels): yes, no

(16)

*Da* *blieben kaum Wünsche offen*. *Das heißt: Wenn man sich das Gerät denn*  
 there stay hardly wishes open this means if one SELF the device DENN  
*leisten konnte*.  
 afford could  
 ‘Hardly any wish was left unfulfilled. That is, if one could DENN afford the device.’\*

Again, we first annotated the items independently; we achieved a percentage of agreement of 65.8% for the first factor ( $\kappa = 0.34$ ) and of 75.4% for the second factor ( $\kappa = 0.45$ ).<sup>13</sup> The items for which we had disagreed were discussed to decide on an annotation; if no clear answer could be found for the first factor, we annotated the item as “indeterminable”.

13 Given the nature of the factors, values of  $\kappa$  of this magnitude are expected (Artstein & Poesio 2008).

### 3.2.3 Results

Tables and Figures 3-4 show the absolute and relative frequencies for the two factors.

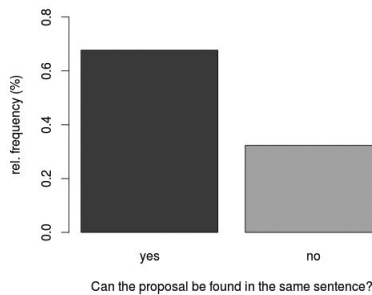
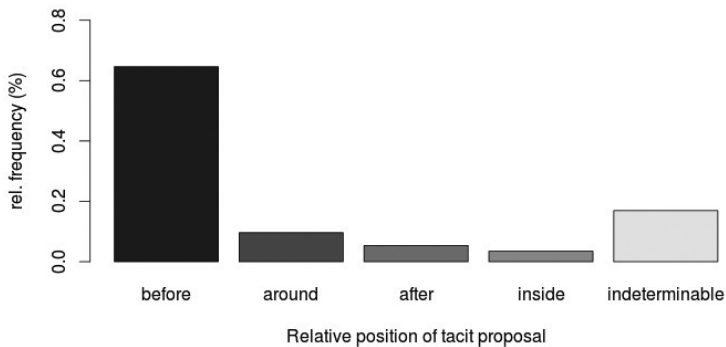
Table 3: Frequencies of REL. POSITION OF PROPOSAL

	BEFORE	AROUND	AFTER	INSIDE	INDETERMINABLE
FREQUENCIES	168 (64.6%)	25 (9.6%)	14 (5.4%)	9 (3.5%)	44 (16.9%)

Table 4: Frequencies of PROPOSAL IN SENTENCE

	YES	NO
FREQUENCIES	176 (67.7%)	84 (32.3%)

Figures 3-4: Rel. frequencies REL. POSITION OF PROPOSAL & PROPOSAL IN SENTENCE



We did not anticipate the existence of the levels “around” and “inside” for REL. POSITION OF PROPOSAL when we formulated the hypothesis in Section 3.2.1. To stay conservative with respect to our hypothesis, we grouped these two levels with

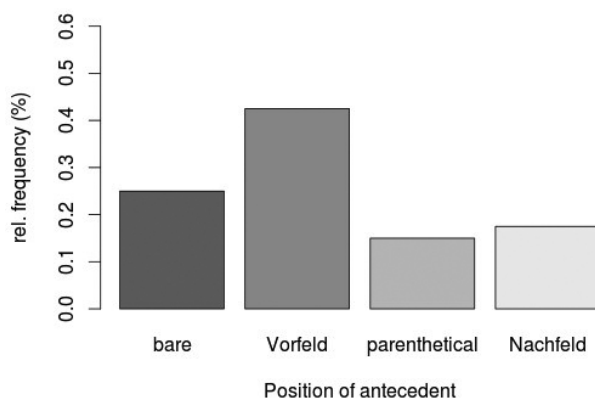
“indeterminable” and “after” for the statistical tests. A  $\chi^2$ -test on this grouping showed that the difference between the level “before” (168 items) and the group containing all other levels (92 items) is statistically significant ( $\chi^2 = 22.22$ ,  $df = 1$ ,  $p < 0.05$ ). Hence, the findings support our hypothesis.

The interaction between the above two factors and POSITION OF ANTECEDENT (Study 1) provides further insights. Items annotated with the levels “around”, “after”, and “inside” (REL. POSITION OF PROPOSAL) were predictably only assigned the level “yes” for PROPOSAL IN SENTENCE. For the level “before”, we find both levels: 128 items “yes” (76.2%) and 40 items “no” (23.8%). At the end of Section 3.1.4., we reported the observation that AWDs seem to be positioned as close to the lexical source material of the tacit proposal as possible. If this is indeed the case, we would expect—for the 40 proposals that precede the AWD (“before”, REL. POSITION OF PROPOSAL), but are not in the same sentence (“no”, PROPOSAL IN SENTENCE)—to find that a majority of these AWDs are either bare or in the Vorfeld (POSITION OF ANTECEDENT). This is borne out, compare Table 5 and Figure 5 to the results in Study 1.

Table 5: Frequencies for POSITION OF ANTECEDENT – items annotated for a preceding proposal that is not in the same sentence as the antecedent

	BARE	VORFELD	PARENTHETICAL	NACHFELD
FREQUENCIES	10 (25%)	17 (42.5%)	6 (15%)	7 (17.5%)

Figure 5: Rel. frequencies for POSITION OF ANTECEDENT – items annotated for a preceding proposal that is not in the same sentence as the antecedent



### 3.3 Summary

Study 1 suggests that some factor (i.e., the presence of conditional *denn*) affects the syntactic placement of AWDs, which preferably occur parenthetically or in the Nachfeld. Our intuitive analysis of the relative position of the lexical source material of the tacit proposals in Study 2 suggests that this material generally precedes the AWD, as is predicted by Condition 2 (see Section 2.2). In addition, the interaction between the factors annotated in the two studies supports the hypothesis that the AWD follows the lexical source material as closely as syntactically possible.

## 4. CLASSES OF TACIT PROPOSALS

During the annotation process for Study 2, we noticed patterns regarding certain classes of tacit proposals. The two main classes of tacit proposals can be distinguished by their origin: proposals made by the speaker and proposals made by others. A third class contains cases that are potentially problematic for which the source of the tacit proposal cannot be identified.

### 4.1 Proposals Made by the Speaker

In the majority of cases, the speaker uses an AWD to qualify her own statements (possibly due to the genre we investigated). We can distinguish the following subclasses.

#### 1) Existence presupposition of a determiner phrase (DP)

The largest individual subclass of tacit proposals are cases like (17), where the speaker uses a definite DP in the consequent, and then inserts an AWD to signal that the existence presupposition triggered by the DP may not hold.

- (17) *Blatter soll besser die Fakten auf den Tisch legen, wenn er denn welche hat.*  
Blatter shall better the facts on the table lay if he DENN some has  
'Blatter had better present the facts if he DENN has any.'\*

In the consequent of (17), the speaker tacitly proposes (or at least assumes) that Blatter possesses "the facts"; otherwise it would not be reasonable to demand that he presents them. The antecedent serves to highlight this assumption and to call it into question.

#### 2) Metacomment

A second class of self-qualifying uses of AWDs is those examples where the speaker adds an antecedent to reflect on word choice, as in (18).

- (18) *Grundsätzlich aber hat Angela Merkel eine Eigenschaft, wenn man es denn*  
basically but has Angela Merkel a trait if one it DENN  
*so nennen will, die mancher Politiker gerne hätte.*  
so call want that many politician gladly had

‘But basically, Angela Merkel has a trait, if you DENN want to call it that, which many politicians would like to have.’\*

Through the use of the AWD in (18), the speaker qualifies her use of the term “trait” in the consequent, which tacitly proposes that the speaker would “call it that”. The AWD also signals the speaker's willingness to change her wording.

### 3) Qualification of a precondition of a modal

For conditionals containing overt modals in the consequent, the antecedent proposition *p* is usually assumed to restrict the modal base of the modal (see Kratzer 2012). AWDs sometimes contain another modal (or attitudinal) proposition that comments on the modal in the consequent, as in (19).

(19)

*Auch bei einem überprüften Lehrplan könnten Lehrer immer noch – wenn sie es*  
also with a checked curriculum could teachers always still if they it  
*denn wollten – ihre ganz eigene Sichtweise des Islams vermitteln.*  
DENN wanted their completely own view of-the Islam convey  
‘After the curriculum has been checked, teachers could still – if they DENN wanted to – convey their very own view of Islam.’\*

The consequent conveys that in view of what the law provides, teachers are allowed to convey their own view of Islam. The speaker, we assume, qualifies her statement by making explicit a tacit assumption suggested by *könnten* (Engl. ‘could’), namely that the teachers actually want to do this. That is, deontic possibility regarding an action only becomes an issue if a wish to perform the action exists.

Since the modal and the dependent infinitive (underlined in (19)) are positioned at the two ends of the German Mittelfeld (see Section 3.1.3), the material providing the tacit proposal usually surrounds a parenthetical antecedent (= level “around”, Study 2).

## 4.2 Proposals Made by Others

The second class of tacit proposals are cases in which the speaker is explicitly arguing against the assumptions that are not her own. In many cases, these assumptions are treated as “established knowledge” by others (i.e., as part of the common ground).

(20)

[Context: American investors are expecting a change in interest rates and have done so for a long time.]

*Einen Überraschungseffekt wie 1987 wird es deshalb nicht geben,*  
a surprise-effect like 1987 will it therefore not give  
*wenn die US-Notenbank denn tatsächlich die Zinswende einläutet.*  
if the US-Federal-Reserve DENN effectively the interest-turn heralds  
‘That is why there will not be a surprise effect if the Federal Reserve DENN does change the interest rates.’\*

The context preceding (20) establishes that expert investors are expecting the Federal Reserve to change the interest rates. By using an AWD, the speaker conveys that she does not share this expectation and that she is skeptical that it will happen.

Sometimes the proposal against which the speaker argues is mentioned explicitly in a complement clause inside the AWD (= level “inside”, Study 2), as in (21).

(21)

[Context: An article about the economic status of various EU accession countries.]

*Wenn es denn zutrifft, dass Ungarn das am höchsten entwickelte Beitrittsland ist,*  
if it DENN is-the-case that Hungary the at-the highest developed accession-country is  
‘If it is DENN the case that Hungary is the best-developed accession country, [...]’\*

Lastly, we got the impression that in some cases, the speaker used an AWD to preemptively argue against something that could be reasonably assumed to be a proposal by somebody else given the topic of discourse.

(22)

[Context: The speaker talks about an artist and his current exhibition.]

*Wenn es denn schon wieder so weit ist, von den Berliner Künstlern der Szene*  
if it DENN already again so far is about the Berlin artists of-the scene  
*der sechziger und siebziger Jahren bereits als Klassiker zu sprechen [...]*  
of-the sixties and seventies years already as classics to speak  
‘If it is DENN already time to call the artists of the Berlin scene in the 60s and 70s  
“classics” [...]’\*

For these examples, no lexical source material could be identified (= level “indeterminable”, Study 2).

### 4.3 Frustration Cases

The frustration cases cannot be clearly classified into one of the two main classes since the question of the origin of the tacit proposal cannot be answered. In these cases, the speaker uses *wenn es denn sein muss* (Engl. ‘if it cannot be helped’) to signal frustration with the state of affairs described in the consequent, and expresses uncertainty that this is the only available option, as in (23).

(23)

[Context: Intimacy and atmosphere in romantic hotels in the US.]

*Das heißt auf Amerikanisch: herzförmige Badewanne, herzförmiges Bett, herzförmige*  
this means on American heart-shaped bath-tub heart-shaped bed heart-shaped  
*Kissen und, wenn’s denn sein muß, auch noch herzförmige Fußabstreifer.*

pillows and if-it den be must also else heart-shaped doormats

‘In the US, this means: a heart-shaped tub, a heart-shaped bed, heart-shaped pillows, and, if it DENN can’t be helped, heart-shaped doormats.’\*

It is unclear whether there has been a previous tacit proposal in these cases or whether the antecedent *wenn es denn sein muss* has become a fixed expression conveying the speaker's frustration about the status quo that does not require a preceding tacit proposal. For reasons of space, this question is left for future work.

## 5. CONCLUSION

We have shown that conditional *denn* is restricted to those types of conditionals which do not commit the speaker to the truth of the antecedent proposition *p* (Condition 1), and contributes that the speaker is skeptical of the truth of *p*. AWDs most frequently occur in contexts where they follow a tacit proposal of *p*, as is required by Condition 2. This explains why AWDs (more so than regular hypothetical antecedents) occur parenthetically or in the Nachfeld. Tacit proposals of *p* can originate with the speaker or with others. In both cases, an AWD signals that the speaker is not or no longer willing to accept *p* as true. Hence, conditional *denn* is a discourse-navigating device (like other discourse particles) that allows the speaker to prevent *p* from becoming common ground.

## References

- ARTSTEIN, Ron/Massimo POESIO (2008) "Inter-coder agreement for computational linguistics." *Computational Linguistics* 34, 555–596.
- BAYER, Josef (2012) "From modal particle to interrogative marker: a study of German *denn*." In: L. Brug et al. (eds), *Functional Heads*. Oxford: Oxford University Press, 13–26.
- BRAUSSE, Ursula (1994) *Lexikalische Funktionen der Synsemantika*. Tübingen: Narr.
- CSIPAK, Eva/Sarah ZOBEL (to appear) "Discourse particle *denn* in the antecedent of conditionals." In: C. Piñón (ed.), *Empirical Issues in Syntax and Semantics 11*. Paris: CSSP. <http://www.cssp.cnrs.fr/eiss11/>
- CONIGLIO, Marco (2011) *Die Syntax der deutschen Modalpartikeln* (studia grammatica 73). Berlin: Akademie Verlag.
- ECKARDT, Regine (2013) "Speaker commentary items." In: *19th ICL papers*, Geneva 20–27 July 2013. Unedited online publication.
- FABRICIUS-HANSEN, Cathrine/Kjell Johan SÆBØ (1983) "Über das Chamäleon *wenn* und seine Umwelt." *Linguistische Berichte* 83, 1–35.
- VON FINTEL, Kai (2008) "What is presupposition accommodation again?" *Philosophical Perspectives* 22/1, 137–170.
- VON FINTEL, Kai (2011) "Discourse particles." In: C. Maienborn et al. (eds), *Handbooks of Linguistics and Communication Science (HSK) 33/2*. Berlin/New York: DeGruyter, 1515–1538.
- HÄUSSLER, Jana (2015) "Wenn *denn* denn lizenziert ist: The German discourse particle *denn* in conditionals." In: E. Brandner et al. (eds), *Webschrift for Josef Bayer*, 83–88. <http://ling.uni-konstanz.de/pages/WebschriftBayer/2015/contents.html>



- HÖHLE, Tilman (1986) “Der Begriff Mittelfeld. Anmerkungen über die Theorie der topologischen Felder.” In: W. Weiss et al. (eds), *Textlinguistik contra Stilistik? – Wortschatz und Wörterbuch – Grammatische oder pragmatische Organisation von Rede? Akten des VII. Internationalen Germanisten-Kongresses, Göttingen 1985. Band 3*. Tübingen: Niemeyer, 329–340.
- KRATZER, Angelika (2012) *Modals and conditionals: New and revised perspectives*. Oxford: Oxford University Press.
- KWON, Min-Jae (2005) *Modalpartikeln und Satzmodus – Untersuchungen zur Syntax, Semantik und Pragmatik der deutschen Modalpartikeln*. LMU Munich.
- MCCREADY, Eric (2006) “Japanese *Yo*: Its semantics and pragmatics.” *Sprache und Datenverarbeitung* 30/1, 25–34.
- POTTS, Chris (2015) “Presupposition and implicature.” In: S. Lappin/C. Fox (eds), *The Handbook of Contemporary Semantic Theory*. Oxford: Wiley-Blackwell, 168–202.
- ROJAS-ESPONDA, Tania (2014) “A discourse model for “überhaupt”.” *Semantics and Pragmatics* 7, 1–45.
- SIMONS, Mandy et al. (2011) “What projects and why.” In: N. Li/D. Lutz (eds), *Proceedings of SALT 20*, 309–327. <http://dx.doi.org/10.3765/salt.v20i0.2584>
- STALNAKER, Robert (1973) “Presupposition.” *Journal of Philosophical Logic* 2, 77–96.
- THURMAIR, Maria (1989) *Modalpartikeln und ihre Kombinationen*. Tübingen: Niemeyer.
- THURMAIR, Maria (1991) “Zum Gebrauch der Modalpartikel “denn” in Fragesätzen. Eine korpusbasierte Untersuchung.” In: E. Klein et al. (eds), *Betriebslinguistik und Linguistikbetrieb*. Akten des 24. Ling. Koll. Bremen, 377–387.
- VOLODINA, Anna (2014) “C4.1 Konditionale Konnektoren.” In: E. Breindl et al. (eds), *Handbuch der deutschen Konnektoren 2. Semantik der deutschen Satzverknüpfers*. Berlin/New York: de Gruyter, 692–786.
- ZIMMERMANN, Malte (2011) “Discourse particles.” In: C. Maienborn et al. (eds), *Handbooks of Linguistics and Communication Science (HSK) 33/2*. Berlin/New York: DeGruyter, 2012–2038.

Corpus: ZEIT corpus of the DWDS online platform, <http://dwds.de/>

### Summary

#### CONDITIONAL ANTECEDENTS CONTAINING THE GERMAN DISCOURSE PARTICLE *DENN*: A CORPUS STUDY

We discuss the semantic contribution and distribution of conditional antecedents containing the discourse particle *denn* (“antecedents with *denn*”, abbreviated as AWD). We propose that AWDs occur only in contexts where (i) the speaker does not believe the antecedent proposition *p* to hold, and (ii) the truth of *p* has been non-explicitly (= tacitly) proposed. To gain a better understanding of (ii), we conduct two

corpus studies. The first study investigates the relative location of AWDs with respect to their consequents. We find that unlike antecedents of regular hypothetical conditionals, AWDs occur significantly more often after the material in the consequent and parenthetically inside this material than before it. In a second study, we investigate the position of the tacit proposal relative to the AWD. We find that it typically precedes the AWD. Both results are in accordance with (ii). We then present a classification of the types of tacit proposals that we find with AWDs: speakers use AWDs to qualify their own statements or to doubt proposals of others, in both cases managing potential updates to the common ground.

**Keywords:** conditional antecedents, German, discourse particles, *denn*, corpus study

#### Povzetek

### NEMŠKI POGOJNI ODVISNIKI Z DISKURZNM ČLENKOM *DENN*: KORPUSNA ANALIZA

Članek obravnava rabo in pomen nemških pogojnih odvisnikov, ki vsebujejo diskurzni členek *denn* (odvisniki z *denn*). V članku predlagamo, da se odvisniki z *denn* rabijo le v tistih kontekstih, v katerih (i) govorec ne verjame v resničnost propozicije *p* in (ii) je resničnostna vrednost propozicije *p* zgolj implicirana. Da bi bolje razumeli slednji kontekst, smo izvedli dve korpusni analizi. V prvi smo se osredinili na položaj odvisnika glede na glavni stavek. Analiza je pokazala, da se odvisniki z *denn* v nasprotju z navadnimi hipotetičnimi odvisniki signifikantno bolj pogosto pojavljajo bodisi za glavnim stavkom bodisi kot vrivek znotraj glavnega stavka kot pa pred glavnim stavkom. V drugi smo proučili relativni položaj konteksta, ki implicira resničnostno vrednost propozicije, glede na položaj odvisnika z *denn*. Podatki pokažejo, da se le-ta pojavlja praviloma pred odvisnikom z *denn*. Oba rezultata sta skladna z (ii). Na koncu predstavimo klasifikacijo kontekstov, ki implicirajo resničnostno vrednost propozicije in omogočajo rabo odvisnikov z *denn*. Govorci uporabljajo odvisnike z *denn*, da kvalificirajo lastne izjave ali da izrazijo dvom do izjav drugih in tako preprečijo, da bi bila propozicija *p* sprejeta kot resnična.

**Ključne besede:** pogojni odvisniki, nemščina, diskurzni členki, *denn*, korpusna analiza





## ON THE (IN)DEFINITENESS OF IMPERSONAL PRONOUNS\*\*

### 1. INTRODUCTION

Although the cross-linguistic variation found for possible uses of impersonal pronouns is quite well-studied, open questions on their semantic analysis remain. One persistent point of contention is whether impersonal pronouns – based on their semantic/pragmatic behaviour – should be analyzed as definite or indefinite expressions. Practically all possible answers to this question have been argued for. They have been analyzed as definite expressions (e.g., Kratzer 1997; Alonso-Ovalle 2002), as indefinite(-like) expressions (e.g., Condoravdi 1989; Moltmann 2012; Malamud 2013), and as “a-definites” (Koenig & Mauner 1999).

The main aim of this paper is to add new empirical facts to the discussion, which, to my mind, tip the scales in favour of Koenig & Mauner’s claim that impersonal pronouns cannot be grouped with either definite or indefinite NPs. The empirical investigation is conducted using the German impersonal pronoun *man*, specifically, its existential use (“existential *man*”).

Like all impersonal pronouns cross-linguistically, German *man* has a generic use, as in (1).

- (1) *Man muss seine Eltern respektieren.* (German)  
MAN has-to his parents respect  
‘One has to respect one’s parents.’

This use occurs exclusively in generic sentences – sentences stating a rule or non-accidental regularity. The existential use of *man* is given in (2).

- (2) *Gestern hat man die Uni angezündet.*  
yesterday has MAN the uni set-on-fire  
≈ ‘Yesterday, someone set the university on fire.’

This use occurs only in episodic sentences – sentences describing a specific situation/eventuality, including accidental generalizations.<sup>1</sup> Unlike the generic use, the

\* sarah.zobel@ds.uni-tuebingen.de

\*\* I thank the audiences of Pronouns@Tübingen 2 and SinFonIJA 8 and two anonymous reviewers for helpful comments and suggestions. All mistakes are my own.

1 Accidental generalizations cannot be used to make predictions about other individuals and are usually expressed by extensional quantification, e.g., with nominal quantifiers: “Every student in

existential use is not uniformly available. English *one*, for instance, lacks this use.<sup>2</sup>

This investigation focuses exclusively on existential *man* since the generic use is inseparably tied to the intensional, quantificational generic operator *Gen*. Since definite and indefinite singular NPs interact with *Gen* in different ways (see Krifka et al. 1995), similarities in the semantic/pragmatic behaviour of *man* and these NPs are always masked by *Gen* (Zobel 2014).

The paper is structured as follows. In Section 2, I present new data on the *discourse anaphoric potential* (DAP) of impersonal pronouns as compared to (in)definite NPs. In Section 3, I show that the DAP of existential *man* is comparable to that of implicit agents of short passives – as Koenig & Mauner (1999) argued for French *on*.<sup>3</sup> Section 4 discusses quantificational variability effects (QVE) with *man*. QVE is seen as the most robust argument for classifying *man* as an indefinite (see Malamud 2013). Section 5 sketches the core idea for a semantic analysis of *man* based on Onea (2013, 2015). Section 6 concludes the paper.

## 2. THE DAP OF (IN)DEFINITE NPS AND EXISTENTIAL *MAN*

The question whether impersonal pronouns are definite or indefinite expressions is **not** discussed in the literature with respect to a single, specific theory of (in)definiteness (see Heim 2011 for a recent overview). The central question, also pursued in this paper, is whether the semantic/pragmatic behaviour of impersonal pronouns is comparable to that of definite or indefinite NPs, or whether they are distinct from either.

The aspect of the semantic/pragmatic behaviour of *man* that I focus on is the *discourse anaphoric potential* (DAP) of existential *man*. That is, for which kinds of anaphoric nominal elements existential *man* can be an antecedent, and conversely, the referents of which nominal elements can be taken up by existential *man*. The DAP of (in)definite NPs is very well studied (see Heim 2011); the data on the DAP of existential *man* is still incomplete (see Cabredo-Hofherr 2010; Malamud 2013 for previous results).

As English *one* lacks an existential use, existential *man* is usually translated as *someone* (see (2)). This translation is inadequate. The scope behaviour of existential *man* is not comparable to that of indefinite pronouns: existential *man*, unlike indefinite pronouns, always takes narrow scope with respect to other quantifiers (Zifonun 2000). Based on the data presented below, I argue that it denotes an indeterminate “group” of individuals (possibly a single person), which I label “X”. Which individuals the speaker means by “X” can, in the right contexts, be inferred.

---

*my class wears a t-shirt*”. Non-accidental generalizations arguably involve intensional quantification (see Greenberg 2007).

2 French *on* (see Koenig & Mauner 1999; Cabredo-Hofherr 2010) and Italian *si* (see Cinque 1988) among other European languages behave like German *man*. Spanish *uno* and Icelandic *maður*, for instance, pattern with English *one* (Cabredo-Hofherr 2010).

3 Koenig & Mauner (1999) do not explicitly restrict their claim to the existential readings of impersonal pronouns, but they only discuss examples of existentially used French *on*.

## 2.1 The DAP of Existential *Man* across Sentence Boundaries

The DAP of (in)definite NPs across sentence boundaries is summarized in (3).

- (3) a) Indefinite NPs can occur discourse initially and can serve as antecedents for strictly anaphoric expressions (i.e., definite NPs and personal pronouns). They cannot take up discourse referents (DRs) that have been previously introduced.
- b) Definite NPs are marked discourse initially. They can serve as antecedents for other strictly anaphoric expressions and can take up DRs that have been previously introduced.

Like indefinite but unlike definite NPs, existential *man* can occur discourse initially.<sup>4</sup> That is, the group X does not have to be previously introduced. In this case, existential *man* is intuitively interpreted similar to an indefinite pronoun.

- (4) *Hast du das gehört? Gestern hat man die Uni angezündet.*  
have you that heard yesterday has MAN the uni set-on-fire  
'Did you hear? Yesterday, X set the university on fire.' (X ≈ someone)

Unlike definite and indefinite NPs, though, X cannot be taken up by 3rd sg. pronouns or arbitrary singular definite descriptions (see Cabredo-Hofherr 2008; Zifonun 2000). None of the expressions in subject position in (5) can refer back to X denoted by *man* in (4).

- (5) #*Der Mann / er / sie hat ein Streichholz in einen Mülleimer geworfen.*  
the man / he / she has a match in a garbage-can thrown  
'The man / he / she threw a match in a garbage can.'

The group X in (4) can, however, be taken up by (i) 3rd pl. personal and demonstrative pronouns with a corporate/bridging reading,<sup>5</sup> (ii) bridging definite NPs (see Schwarz 2009), and (iii) existential *man*. This is illustrated in (6) (= i & ii) and (7) (= iii), which can both continue (4).<sup>6</sup>

---

4 The judgments given on the German data are based on my own native speaker judgments and judgments of three native speaker informants (of different varieties) on constructed and naturally occurring examples. Experimental evidence is still to be gathered.

5 Koenig & Mauner (1999) call the corporate reading of 3rd person plural personal pronouns their "indefinite use". I prefer the more theoretically neutral term "corporate/bridging reading" since it is not clear that the pronouns in this use are indeed indefinite (see Malamud 2013; Grosz 2016).

6 Malamud (2013) argues that the cases illustrated in (6) are impossible for German *man*. Her claim is based on the result of a corpus study: her sample (100 occurrences of *man*) did not contain any examples of this kind. However, 100 tokens of *man* are too few to establish all potential continuations for *man*, and existential *man* is itself quite rare.

(6)

*Sie / die / die Brandstifter haben ein Streichholz in einen Mülleimer geworfen.*  
they / d-pron / the arsonists have a match in a garbage-can thrown  
'They / they / the arsonists threw a match in a garbage can.'

(7) *Man hat ein Streichholz in einen Mülleimer geworfen.*

MAN has a match in a garbage-can thrown  
'X' threw a match in a garbage can.' (X' → X of (4))

The expressions in (6), which intuitively refer to X in (4), are not strictly anaphoric to it. The perceived "reference sharing", I argue, is the result of inference. This is supported by the different number specifications of *man* (sg.) and *sie/die/die Brandstifter* (pl.); strictly anaphoric expressions (e.g., co-referring personal pronouns and definite NPs) always agree with their antecedents in person and number.<sup>7</sup>

Subsequent occurrences of existential *man* may but do not have to refer back to the group of individuals given by a preceding occurrence of existential *man*. Example (9) can continue (4), just like (8).

(8) *Man sucht noch nach den Brandstiftern.*

MAN searches still for the arsonists  
'X'' is still looking for the arsonists.' (X'' → the police)

Since the group understood for existential *man* in (4), the arsonists, are explicitly referred to in (8), another group of individuals has to be inferred for *man* in (8). World knowledge suggests that the people looking for the arsonists are most plausibly the police.

Lastly, existential *man* can be interpreted as denoting previously introduced DRs, as in (9). This possibility is only shared by definite NPs.

(9)

*Eine Gruppe von Studenten ist für ihren Vandalismus bekannt. Gestern hat man zum*  
a group of students is for their vandalism known yesterday has MAN for  
*Beispiel die Uni angezündet.*  
example the uni set-on-fire  
'A group of students is known for their vandalism. For example, yesterday X set the  
university on fire.' (X → the group of students)

In (9), *zum Beispiel* (Engl. 'for example') signals that the second sentence takes up the subject matter of the preceding sentence. Hence, the group of students introduced

---

7 One counterexample to this rule are split antecedents, see (i). Example (4) does not provide the right setup to argue that *sie* (Engl. 'they') has a split antecedent, though.  
(i) *Every boy/Peter has asked some girls/Mary if they could go out on a date.* (see Büring 2011: 988)

in the first sentence is a plausible candidate for the agents of the second sentence (= X). Crucially, the speaker in (9) does not explicitly claim that the group of students is responsible for setting the university on fire, which would be the case if she had used the strictly anaphoric 3rd pl. personal pronoun *sie* (Engl. ‘they’). This, I argue, is a result of determining the specification of X via inference.

In general, highly topical or salient DRs can be inferred as “referents” of *man* – provided that the discourse relations that link the utterances, as in (9), do not discourage this inference (see Asher & Lascardes 2003 on discourse relations).

## 2.2 The DAP of Existential *Man* Sentence Internally

Existential *man* can “co-refer” with (i) the 3rd singular reflexive *sich*, (ii) PRO, and (iii) existential *man*, but, again, not with personal pronouns or definite NPs (Cabredo-Hofherr 2008; Zifonun 2000).<sup>8</sup> For (in)definite NPs, no such restrictions obtain. Possibilities (i) and (ii) are illustrated in (10).

- (10) *Man hat sich entschlossen PRO zu bleiben.*  
 MAN has SELF decided PRO to stay  
 ‘X decided to stay.’ ( = X decided that X stays)

The possibility of reflexivization and control for existential *man* is not a counter-argument against the claim that *man* cannot co-refer with strictly anaphoric expressions. Here I follow Chierchia (1995) and Landau (2010) in assuming that reflexivization and control do not involve co-reference.

Multiple occurrences of existential *man* in multi-clausal sentences (possibility (iii) above) can again be read either as referring to the same group or a larger group of individuals, or as referring to two (not necessarily overlapping) groups, as in (11). The former reading is preferred.

- (11) *Man hat hoffnungsvoll gefragt, ob man sich morgen trifft.*  
 MAN has hopeful asked whether MAN SELF tomorrow meets  
 ‘X asked hopeful whether X’ are meeting up tomorrow.’

## 2.3 Interim Summary

From Sections 2.1 and 2.2, I conclude that existential *man* and (in)definite NPs differ as follows:

- Indefinite NPs always introduce new DRs and, hence, cannot refer to DRs that were previously introduced.
- Definite NPs (almost) always refer to DRs that were previously introduced.

8 For existential *man*, the suppletive forms *einen* (acc.) and *einem* (dat.)—inflected forms of the indefinite pronoun *einer* (Engl. ‘someone’)—are not available, and existential *man* cannot “co-refer” with possessives; both is possible for generically used *man* (see Kratzer 1997). An investigation of this issue has to be left for future research.



- Existential *man* never introduces new DRs that could be referred to by anaphoric expressions, and cannot refer to DRs that were previously introduced.

As Condoravdi (1989), Moltmann (2012), and Malamud (2013), among others, argue, *man* contributes a free variable that is, in the generic use of *man*, unselectively bound by *Gen*. If this idea is to be extended to existential *man*, one has to find a way to distinguish *variables*, which are needed to model quantification, from *DRs*, which are needed to model anaphoric relations, and find a way to connect these two appropriately. The core idea for an analysis with these features is sketched in Section 5.

### 3. EXISTENTIAL *MAN* AND IMPLICIT AGENTS

The data on the DAP of existential *man* given in Section 2 mirrors the DAP of German implicit agents of short passives (IAP), which are “strong implicit arguments” (Landau 2010). IAPs in German can occur discourse initially, as in (12a), but cannot be taken up in a subsequent sentence by strictly anaphoric expressions, as in (12b).

- (12) a) *Der Uni wurde IAP abgefackelt.*  
 the uni was IAP set-on-fire  
 ‘Someone set the university on fire.’  
 b) #*Der Mann / er / sie hat Benzin verwendet.*  
 the man / he / she has gas used  
 ‘The man / he / she used gas.’

The IAP in (12a) can be picked up by bridging definites and corporate/bridging pronouns, as in (13), as well as another IAP, as in (14).

- (13) *Die Täter / sie haben Benzin verwendet.*  
 the offenders / they have gas used  
 ‘The offenders / they used gas.’  
 (14) *Es wurde IAP Benzin verwendet.*  
 it was IAP gas used  
 ‘They used gas.’

Sentence-internally, IAPs can control PRO and “co-refer” with reflexive *sich*, as in (15), see Alexiadou & Müller 2015.

- (15) a) *Es wurde IAP beschlossen PRO Benzin zu verwenden.*  
 it was IAP decided PRO gas to use  
 (≈ Someone/X decided to use gas.)  
 b) *Hier wurde IAP sich nicht geprügelt.*  
 here was IAP SELF not hit  
 ≈ ‘No one hit each other here.’

In sum, this strong parallel in the DAP of existential *man* and IAPs suggests that they should indeed be analyzed similarly (*pace* Malamud 2013). Koenig & Mauner (1999) observe the same characteristics for French IAPs and the impersonal pronoun *on*. They introduce the notion of “a-definites” in (16) to refer to expressions with this DAP (compare to Section 2.3).

- (16) A-definites are expressions that are “inert in discourse”: they cannot serve as the anchor of an anaphoric element—unless the perceived anaphoricity is the result of lexical or inferential processes. (Koenig & Mauner 1999: 213, 220ff)

#### 4. IMPERSONAL PRONOUNS AND QVE

As the main argument for classifying impersonal pronouns as indefinite-like expressions, the observation is usually given that impersonal pronouns show “classical” quantificational variability effects (QVE) with adverbs of quantification like *often*, *usually*, or *seldom* (see Malamud 2013), which do not occur with singular definite expressions (see Hinterwimmer 2008). Example (17) illustrates classical QVE with German *man* and English *one*.<sup>9</sup>

- (17) *An dieser Fakultät ist man normalerweise klug.* (≈ Malamud 2013: 14)  
 at this department is MAN usually smart  
 ‘In this department, one is usually smart.’  
 (≈<sub>QVE</sub> Most members of this department are smart.)

In addition, Malamud (2013: 26) observes that English IAPs show only QVE-like effects with *for the most part*, as in (18). Such QVE-like effects have been reported to occur only with plural definite expressions (Nakanishi & Romero 2004).<sup>10</sup>

- (18) *In Spain, Michael Jackson is for the most part admired.* (Malamud 2013: 21)  
 (≈<sub>QVE</sub> Most Spaniards admire Michael Jackson.)

Hence, the availability of QVE vs. QVE(-like) effects apparently differentiates impersonal pronouns from implicit agents. That is, the result of Section 3 that existential *man* and IAPs show parallel behaviour seems to be incorrect.

However, this conclusion is premature. Firstly, there are cases of classical QVE with German IAPs: (19) can be interpreted as stating that the majority of implicit agents (i.e., doctors/researchers) assume the given list of reasons. That is, *üblicherweise* (Engl. ‘usually’) quantifies over implicit agents.

9 All sentences for which I claim an availability of QVE or QVE(-like) effects also have a second interpretation for which the adverb of quantification quantifies over times, situations, or parts.

10 Alexiadou & Müller (2015) also observe QVE-like effects for German IAPs.

- (19) *Zu den Ursachen des Herzinfarkts werden üblicherweise [...] gezählt.*  
 to the reasons of-the heart-attack will usually [...] counted  
 ( $\approx_{\text{QVE}}$  Most doctors/researchers count [...] among the reasons for a heart attack.)<sup>11</sup>

Secondly, existential *man* also shows QVE-like effects with *größtenteils* (Engl. ‘for the most part’).

- (20) [Context: Newspaper article on a theatre performance in a local park.]  
*Man war größtenteils in legerer Sommerkleidung gekommen.*  
 MAN was for-the-most-part in casual summer-dress come  
 ‘For the most part, X had appeared in casual summer dress.’<sup>12</sup> (X → the audience)  
 ( $\approx_{\text{QVE}}$  Most people in the audience had appeared in casual summer dress.)

Together, (17) and (20) would imply that *man* has to be classified and analyzed as both indefinite and definite, which is an undesirable result. I believe that the possibilities regarding QVE vs. QVE-like effects vary with the uses of *man*. For reasons of space, further details have to be left for another occasion. The upshot is that the possibility of QVE with *man* is not air-tight evidence that *man* is an indefinite(-like) expression.

## 5. SKETCHING A FORMAL ANALYSIS FOR *MAN*

To capture the DAP of existential *man*, we need a formal system that can distinguish between variables and DRs (see Section 2.3). While the formal system proposed in (Onea 2013, 2015) is not explicitly designed to do this, it can be extended to capture this distinction.

In Onea’s (2013, 2015) system<sup>13</sup>, all lexical entries take assignment functions as arguments. “Referential expressions” (i.e., (in)definite NPs, proper names, pronouns) place constraints on these assignments. A proper name like *Peter*, as in (21), contributes the value  $h(i)$  (= an individual) returned by the assignment argument  $h$  for its index  $i$ , provided that the restriction on the assignment,  $h(i) = \textit{Peter}$ , is met.

- (21)  $[[ \textit{Peter}_i ]]^w = \lambda h[h(i) = \textit{Peter}]. h(i)$

A sentence like *Peter laughs* is assigned the denotation in (22), which is true for an assignment  $h$  iff  $h(i) = \textit{Peter}$  and  $h(i)$  is laughing in  $w$ . The restriction on compatible assignments contributed by *Peter* in (21) is inherited by the full sentence.

- (22)  $[[ \textit{Peter}_i \textit{ laughs} ]]^w = \lambda h[h(i) = \textit{Peter}]. \textit{laugh}(h(i))(w)$

11 <http://www.zeit.de/1966/51/wie-entsteht-der-herzinfarkt> (last accessed: Aug 25, 2016)

12 [http://www.nwzonline.de/cloppenburg/kultur/maerchen-in-schrillen-kostuemen\\_a\\_30,0,1014137318.html](http://www.nwzonline.de/cloppenburg/kultur/maerchen-in-schrillen-kostuemen_a_30,0,1014137318.html) (last accessed: Aug 25, 2016)

13 For reasons of space, I cannot discuss the system in more detail. I refer the interested reader to the original works.

In this system, quantification and binding both utilize the assignment arguments. Quantifiers quantify over sets of assignments; pronouns denote the output that their assignment argument provides for the index that they bear, as in (23).

$$(23) \quad [[\text{pron}_i]]^w = \lambda h. h(i)$$

Onea’s system is only designed to handle inter-sentential binding and anaphora. To model cross-sentential anaphora, I extend it by a parameter  $G$ , which records the active DRs.

$G$  is a set of assignment functions. At the start of the conversation,  $G$  equals the set of all assignment functions  $A$ . Each subsequent sentence reduces this set. For instance, the denotation of *Peter laughs* in (22) removes all assignment functions in  $G$  that do not output Peter for the index  $i$  or for which the individual returned for  $i$  does not laugh in  $w$ .

For *man*, I assume that it has the same denotation as anaphoric pronouns, as in (23), which is equivalent to assuming that *man* contributes a free variable in more familiar static systems.

$$(24) \quad [[\text{man}_i]]^{w,G} = \lambda h. h(i)$$

To ensure that existential *man* does not access or restrict the set  $G$  (i.e., does not access or contribute a DR), we need to assume that it is bound by a selective variant of existential closure at the VP level (Onea 2015). The denotation of *Man hat gelacht* (Engl.  $\approx$  ‘Someone laughed’) is as in (25).

$$(25) \quad [[\exists_i \text{man}_i \text{ hat gelacht}]]^{w,G} = \lambda h. \exists g [g =_i h \exists L(g(i))]$$

The selective existential closure operator  $\exists_i$  in (25) introduces existential quantification over assignments  $g$  that are identical to the assignment argument  $h$  except for the output for the index  $i$  ( $g =_i h$ ). Since only the restrictions placed on  $h$  will constrain  $G$ , this means that any content that is predicated of  $g(i)$  will not access or add restrictions to  $G$ . Conceptually, this ensures that existential *man* cannot refer to existing DRs or introduce new DRs, as desired.

The generic use of *man* is captured by assuming that *man* is bound by *Gen* at the sentence level (see Condoravdi 1989; Moltmann 2012; Malamud 2013), and QVE with *man* can be modeled by assuming that it is bound by an adverb of quantification (see Malamud 2013). For reasons of space, I cannot present this proposal and its implications in any more detail.

## 6. CONCLUSION

I have shown that the DAP of existential *man* differs from that of indefinite and definite NPs, but is parallel to that of German IAPs, which can be classified as “a-definites” following Koenig & Maurer (1999). Furthermore, I showed that using the availability

of QVE as an argument for the claim that impersonal pronouns are indefinite is not as straightforward as has been previously claimed. Lastly, I sketched a formal analysis that can capture the DAP of existential *man* outlined in this paper.

## References

- ALEXIADOU, Artemis/Gereon MÜLLER (2015) “Passive Operations: The External Argument.” Poster presented at *Sinn und Bedeutung 20*, University of Tübingen.
- ALONSO-OVALLE, Luis (2002) “Arbitrary Pronouns Are Not That Indefinite.” In: C. Beyssade et al. (eds), *Romance Languages and Linguistic Theory 2000*. John Benjamins, 1–15.
- ASHER, Nicholas/Alex LASCARIDES (2003) *Logics of Conversation*. Cambridge: Cambridge University Press.
- BÜRING, Daniel (2011) “Pronouns.” In: C. Maienborn et al. (eds), *Handbooks of Linguistics and Communication Science (HSK) 33/2*. Berlin/New York: DeGruyter, 971–996
- CABREDO-HOFHERR, Patricia (2008) “Les pronoms impersonnels humains - syntaxe et interprétation.” *Modèles linguistiques tome XXIX-1/57*, 35–56.
- CABREDO-HOFHERR, Patricia (2010) “Binding properties of impersonal human pronouns in generic and episodic contexts.” Handout presented at the *Workshop on impersonal pronouns*, May 28, Paris. <http://archive.sfl.cnrs.fr/sites/sfl/IMG/pdf/impersMay2010CabredoHofherrManOnLa.pdf>
- CHIERCHIA, Gennaro (1995) “The Variability of Impersonal Subjects.” In: E. Bach et al. (eds), *Quantification in Natural Languages*. Springer, 107–143.
- CINQUE, Guglielmo (1988) “On *si* constructions and the theory of arb.” *Linguistic Inquiry* 19, 521–582.
- CONDORAVDI, Cleo (1989) “Indefinite and generic pronouns.” In: E. Jane Fee/K. Hunt (eds), *Proceedings of the Eighth West Coast Conference on Formal Linguistics*. Stanford, CA: Center for the Study of Language and Information, 71–84.
- GREENBERG, Yael (2007) “Exceptions to generics: Where vagueness, context dependence and modality interact.” *Journal of Semantics* 24, 131–167.
- GROSZ, Patrick (2016) *Bridging uses of demonstrative pronouns in German*. Ms. University of Oslo.
- HEIM, Irene (2011) “Definiteness and indefiniteness.” In: K. von Heusinger/C. Maienborn/P. Portner (eds), *Semantics*, Volume 2. Berlin/Boston: De Gruyter Mouton, 996–1025.
- HINTERWIMMER, Stefan (2008) *Q-Adverbs as Selective Binders: The Quantificational Variability of Free Relatives and Definite DPs*. Berlin/New York: De Gruyter.
- KOENIG, Jean-Pierre/Gail MAUNER (1999) “A-definites and the Discourse Status of Implicit Arguments.” *Journal of Semantics* 16, 207–236.
- KRATZER, Angelika (1997) “German impersonal pronouns and logophoricity.” Presentation at *Sinn und Bedeutung II*. Berlin, Germany. [semanticsarchive.net](http://semanticsarchive.net).
- KRIFKA, Manfred et al. (1995) “Genericity: an introduction.” In: G. Carlson/F.J. Pelletier (eds), *The Generic Book*. Chicago: University of Chicago Press, 1–124.

- LANDAU, Idan (2010) "The explicit syntax of implicit arguments." *Linguistic Inquiry* 41, 357–388.
- MALAMUD, Sophia A (2013) "(In)definiteness driven typology of arbitrary items." *Lingua* 126, 1–31.
- MOLTMANN, Friederike (2012) "Two kinds of first-person-oriented content." *Synthese* 184, 157–177.
- NAKANISHI, Kimiko/Maribel ROMERO (2004) "Two constructions with *most* and their semantic properties." In: K. Moulton/M. Wolf (eds), *Proceedings of the North East Linguistic Society* 34/2. Amherst: University of Massachusetts at Amherst, 453–468.
- ONEA, Edgar (2013) "Indefinite donkeys on islands." In: Todd Snider, *Proceedings of SALT* 23, University of California, Santa Cruz, 493–513. <http://dx.doi.org/10.3765/salt.v23i0.2666>.
- ONEA, Edgar (2015) "Why indefinites can escape scope islands." *Linguistics and Philosophy* 38, 237–267 .
- SCHWARZ, Florian (2009) *Two Types of Definites in Natural Language*. Amherst: University of Massachusetts at Amherst.
- ZIFONUN, Gisela (2000) "Morphosyntax und Semantik des Pronomens "man"." *Deutsche Sprache* 28, 232–253.
- ZOBEL, Sarah (2014) *Impersonally Interpreted Personal Pronouns*. University of Göttingen. <http://hdl.handle.net/11858/00-1735-0000-0023-991B-2>

#### Summary

#### ON THE (IN)DEFINITENESS OF IMPERSONAL PRONOUNS

This paper addresses the question whether impersonal pronouns should be analyzed as indefinite or definite expressions based on their *discourse anaphoric potential*. I present new data that support the claim that impersonal pronouns should be analyzed as neither (see Koenig & Mauner 1999). I sketch a formal analysis that captures this behaviour. Furthermore, I show that the availability of quantificational variability effects for impersonal pronouns is not foolproof evidence for their indefiniteness as is usually assumed in the literature (see Malamud 2013).

**Keywords:** impersonal pronouns, German, *man*, existential use, discourse anaphoric potential

#### Povzetek

#### O (NE)DOLOČNOSTI NEOSEBNIH ZAIMKOV

V članku se ukvarjamo z vprašanjem, ali naj neosebne zaimke glede na njihov *diskurzivni anaforični potencial* analiziramo kot nedoločne ali določne izraze. Predstavimo nove podatke, ki govorijo v prid trditvi, da neosebni zaimki ne spadajo v nobeno od obeh kategorij (glej Koenig/Mauner 1999). Predlagamo formalno analizo, ki odraža to

njihovo vedenje. Poleg tega pokažemo, da prisotnost učinkov kvantifikacijske variabilnosti v primeru neosebni zaimkov ni neizpodbiten dokaz za njihovo nedoločnost, kot se običajno predpostavlja v literaturi (cf. Malamud 2013).

**Ključne besede:** neosebni zaimki, nemščina, *man*, bivanjska raba, diskurzivni anaforični potencial

LINGUISTICA LVI

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

Branka Kalenić Ramšak

Dekanja Filozofske fakultete – Doyen de la Faculté des Lettres

Glavna in odgovorna urednica – Rédactrice en chef

Martina Ožbot

Številko LVI uredil – Numéro LVI dirigé par

Gasper Ilc, Frančiška Lipovšek, Tatjana Marvin, Andrej Stopar

Tajnica redakcije – Secrétaire de rédaction

Metka Šorli

Dopise nasloviti 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: 300 izvodov – Tirage : 300 exemplaires

Računalniški prelom – Mise en page

Aleš Cimprič

Tisk – Impression

Birografika BORI, d. o. o.

Linhartova cesta 1, 1000 Ljubljana

Cena: 17 €





