



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
Tokens	3	2	1	1	2	11	13	33
# harmonic	/	/	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

	i	o
feminine	70	113
masculine	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

	i	o
alveolar	56	127
labial	22	40
palatal	13	13
velar	11	10
laryngeal	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

preceding vowel		i	o
FRONT	E	41	95
	I	6	23
	Ö	8	5
	Ü	1	8
TOTAL		56	131
BACK	A	28	51
	U	14	11
	O	10	2
TOTAL		52	64

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*. 2nd 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 17th 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