2024. Proceedings of the Workshop on Turkic and Languages in Contact with Turkic 9. 121–134.

/e/ lowering in Turkish

Mehmet Can Dadan, Uğur Erken & Stefano Canalis*

Abstract. Descriptions of the phonological system of Turkish concur that Turkish /e/ is lowered in closed syllables ending in a sonorant consonant. Nevertheless, despite the consensus on the general traits of this lowering process, there is some disagreement about certain aspects of it, and these conflicting claims are generally unsupported by instrumental data. Furthermore, lexical exceptions to the rule exist, but discussions of them are limited. Here we investigate acoustically this vowel lowering process and examine some of its lexical exceptions, focusing on cases of unexpected lowering before /z/ codas. We argue that this lowering process is no longer purely allophonic, and exhibits properties typical of the early phases of so-called 'lexical' phonological splits, with signs of phonologization of a marginal /e/ ∼ /æ/ contrast.

Keywords. Turkish phonetics; vowel lowering; phonologization; first formant; negative aorist suffix

1. Introduction. Descriptions of the phonological system of Turkish (see e.g. Göksel & Kerslake 2005; Erguvanlı Taylan 2015; Kornfilt 1997) concur that Turkish /e/ (and to a lesser extent /ø/) has lowered allophones. For instance, Göksel & Kerslake (2005) state that while [e] occurs in most environments, [ɛ] occurs in word-final open syllables and [æ] occurs in closed syllables where the coda is /ɾ, l, m, n/.[1] All three sounds can be obverved in a single word (1).

(1) a. *gezegen-de* [ɟezeɟændɛ] 'on the planet', *perende* [perændɛ] 'somersault'

Alternations between [e] and [æ] caused by resyllabification are systematic (2) (Göksel & Kerslake 2005).

* We would like to thank two anonymous reviewers, Didar Akar, Umur Karakaş, Utku Türk, and the audience at TU+9 for their comments and suggestions, and the data annotators for their time and effort. Authors: Mehmet Can Dadan, Boğaziçi University (mehmet.dadan@bogazici.edu.tr), Uğur Erken, Università di Siena (u.erken@student.unisi.it) & Stefano Canalis, Boğaziçi University (stefano.canalis@bogazici.edu.tr).

[1] Hereafter, we assume that the environment triggering lowering is (lexical exceptions such as *pekmez* 'molasses' [pekmæz] aside) a [+sonorant] coda consonant. The phonemic inventory of Turkish also includes /j/, which does not trigger lowering (cf. *bey* 'mister' [bej]), seemingly contradicting the claim that sonorant consonants condition the process. However, the distribution and phonological behavior of the front glide suggest that it does not function as a phonological sonorant in Turkish (Gopal & Nichols 2022; Canalis et al. 2024). A more phonetically oriented explanation might attribute the blocking of vowel lowering before /j/ to the high tongue body position required for its articulation, which could conflict with the low tongue body position needed for the lowering of an immediately preceding /e/. Yet, this account would be difficult to reconcile with the properties of Turkish laterals. Turkish laterals are palatalized when their tautosyllabic vowel is front (Clements & Sezer 1982), and acoustic measurements show that their palatalization is strong (Gopal & Nichols 2022: 42). This means that in words such as *tünel* 'tunnel' [tynælʲ] the final vowel is followed by a consonant with a tongue body position comparable to (and hence roughly as high as) that of a palatal semivowel. Nevertheless, both Gopal & Nichols (2022) and our own data show that vowel lowering is neither blocked nor reduced before /lʲ/.

(2) a. *gel* [ɟæl] 'come!', *gel-di-m* [ɟældim] 'I came', *gel-ir* [ɟelir] 's/he comes'
b. *ev-ler* 'houses' [evlær], *ev-ler-de* [evlærdɛ] 'in (the) houses', *ev-ler-im* [evlerim] 'my houses'

However, in spite of its regularity in a huge number of words, lowering to [æ] — which is the focus of this paper — both under- and over-applies (with some amount of inter-speaker and possibly intra-speaker variability). As for under-application, in a small number of words some speakers realise /e/ before /l/, /m/ and /n/ as either [e] or [æ]. Hence, the words *elbise* 'dress', *kendi* 'self', *bel* 'waist' or *hem* 'both' can occur with either [e] or [æ] (Göksel & Kerslake 2005: 11); for some younger speakers the form with [e] seems to be the only possible one, creating near-minimal pairs (*bel* [bel] vs. *belge* [bælɟɛ] 'document', *kendi* [cendi] vs. *kent-li* [cæntli] 'urban', and so on). Along with the age factor, geographical variation plays a role too: "in large parts of Anatolia, /e/ [...] is preserved as a phoneme beside unrounded mid front /ä/, allowing the existence of minimal pairs, such as *el* 'stranger', 'country' and *äl* 'hand'" (Brendemoen 2021). Speakers of these dialects may carry this pronounciation in their use of standard Turkish.

Additionally, the allophone [æ] may have negative socio-linguistic connotations and is sometimes actively avoided by certain speakers. We have not been able to find any study addressing these socio-linguistic connotations, so the following observations are based on impressionistic observations. Especially in high-frequency words such as *ben* 'I,' *sen* 'you,' and *en* 'most,' some speakers tend to use a higher vowel instead of [æ]. When asked why, they reported associating the use of [æ] with a low socio-economic class, a low education level, and a stereotypical 'shallow young urban woman' usage. At the same time, *absence* of lowering in words such as *el* 'stranger' may connote a regional, non-standard pronunciation, as noted in the previous paragraph.

The lowering rule over-applies as well: in an even more limited number of words, /e/ seems to be lowered to [æ] when the coda is [z]. Gopal & Nichols (2022) report lowering in the final syllables of the words *merkez* 'centre', *pekmez* 'molasses' and the front-harmonized realization of the negative aorist suffix *-mAz*, but they do not expand on their finding. Both under- and over-application make the occurrence and non-occurrence of [æ] not fully predictable.

Beyond these general traits, not much else can be found in the literature on this process. Detailed acoustic analyses are lacking, apart from the rich but still unpublished survey reported in Gopal & Nichols (2022), and discussions of lexical exceptions are typically limited to short lists of examples. In this paper, we provide an acoustic description of this process and briefly discuss some lexical exceptions to the otherwise allophonic lowering rule.

The structure of this paper is as follows: in section 2 we outline our research questions and goals. In section 3 we describe the design of our acoustic survey and the corpus frequency data we used. We report and discuss our findings in section 4, and present our conclusions in section 5.

2. Research Questions. Given the paucity of instrumental phonetic data on this process and the preliminary nature of our study, most of our goals are descriptive and concern the acoustic realization of /e/ before sonorant codas, as well as comparisons to its realization in open syllables and before obstruent codas.

First, we investigate whether the allophone occurring before sonorant codas is realized as [ɛ], as stated by Kornfilt (1997), or as [æ] as stated by Göksel & Kerslake (2005). Second, we

examine whether it forms a phonetic continuum with the realizations of /e/ in the other environments or, as reported by Gopal & Nichols (2022), they form two distinct, fundamentally non-overlapping sets. Third, we examine the environment triggering lowering, with the aim of verifying whether all and only sonorant consonants cause lowering, and by roughly the same amount. Fourth, we want to see if our data replicate Gopal & Nichols's (2022) finding that lowering also occurs before a /z/ coda in a handful of morphemes; fifth, if lowering is as pronounced there as before sonorant consonants. Finally, although it is not the primary focus of this paper, we examine the height of /e/ in word-final open syllables (i.e., the ones where [ɛ] occurs according to Göksel & Kerslake 2005) and word-final closed syllables ending in an obstruent.

In addition to these acoustic investigations, we also briefly address the role of lexical frequency in the mechanisms that may have led to cases of overapplication of the rule (i.e., lowering before /z/). The occurrence of [æ] in a handful of lexical items, as well as its non-occurrence before a sonorant coda in a few other words, means that [e] and [æ] have been acquiring a contrastive status, albeit quantitatively very limited (exceptions occur in a tiny fraction of the closed syllables in which /e/ occurs), in a restricted phonological environment (only in closed syllables), and variable across speakers (not every speaker has all the lexical exceptions mentioned above). This embryonic contrastivity suggests that the Turkish phoneme /e/ is undergoing a phonological split, some occurrences of /e/ no longer being predictable allophones, and it is possibly at the early stage of a phonologization process leading to a phonological split and the emergence of a new phoneme /æ/.

Such a sound change would be a case of 'lexical' phonological split (that is, a division of one phoneme into two that has some lexical exceptions to any phonological conditioning, and therefore cannot be fully accounted for in purely phonological terms). Being at least partially lexeme-specific, lexical splits differ from fully regular, Neogrammarian sound changes, and have been argued (Labov 1994: 333) to show properties that typically are unattested in the latter, and which include among others:

1. "[a] complex set of phonetic conditions that predicts the great majority of lexical assignments, but not all".
2. "[a] number of phonetic conditions that specify particular words", hence the lexical exceptions (Labov 1994: 335 clarifies that he refers to rules that may have numerous lexical exceptions, but whose environment is not entirely phonologically arbitrary).
3. "grammatical conditions" involving morphological categories and morphological boundaries
4. "[v]ariable behavior of derived words"
5. "[i]rregular patterns of proper names and learned words depending on the age of acquisition"
6. "[a] strong effect of word frequency", with the more frequent words favored in the change
7. phonologically "[u]nmotivated lexical exceptions"

As for point 5, we did not include these items in our experiment and previous studies do not address them; therefore, we will not comment on them. On the other hand, we believe all the properties mentioned in the other six points can be found in Turkish /e/ lowering. We discuss each of them in the following sections.

2.1. COMPLEX SET OF PHONETIC CONDITIONS. Properties 1, 2 and 3 can be observed in the examples provided above. As for the first, the rule /e/ → [aɐ] __$\mathrm{C}_{[+\text{sonorant}]}$. correctly predicts almost all the occurrences of [e] as well as those of [aɐ] (i.e., the rule usually predicts when it should or should not apply), but not all. Additionally, its environment is the intersection of two separate dimensions (syllable type and sonority), which, as Gopal & Nichols (2022) observe, commonly affect vowel height independently, but rarely are both needed in vowel lowering processes. Typologically, closed syllables favour vowel lowering, but usually independently of the manner or place of the coda consonant. Symmetrically, although sonorant-triggered height effects are common, they are rarely associated with specific syllable types. Furthermore, sonorants often do not function as a uniform class in relation to vowel lowering. Gopal & Nichols (2022: 22-23) note that while rhotics exhibit articulatory and acoustic properties strongly favoring vowel lowering, nasals are typologically associated with both phonetic raising and lowering, and palatalization in laterals, which consistently occurs in Turkish coda laterals preceded by a front vowel (cf. footnote 1), should cause an F1 decrease, and hence perceptual *raising*.

2.2. PHONETIC CONDITIONS THAT SPECIFY PARTICULAR WORDS. Cases of overapplication of the lowering rule are extremely limited (reportedly, the suffix *-mez*, the roots *pekmez* and *merkez*), but they share a common environment: the following consonant is /z/. Thus, we have an additional conditioning that has a lot of exceptions (the overwhelming majority of syllables containing /e/ and ending in /z/ do not display lowering), but is not phonetically arbitrary.

2.3. CONDITIONS INVOLVING MORPHOLOGICAL CATEGORIES AND BOUNDARIES. The presence of property 3 in the lowering process is less *prima facie* obvious, at least if we only consider the data presented so far. However, with regard to the role of morphological conditioning, the vowel lowering process is not completely oblivious to non-phonological information, as it is sensitive to word boundaries; while resyllabification inside a word regularly creates alternations, as shown in (2), it does not across words (e.g., [askæ.r#itʃti] *asker içti* 'the soldier drank' rather than *[askɐ.r#itʃti]. In Stratal OT terms (Kiparsky 2015), Turkish vowel lowering is a word-level process.

A consequence of this restriction is that lowering in a word-final syllable indirectly signals the presence of a word boundary after the immediately following consonant, giving the vowels [e] and [æ] a quasi-contrastive status in pairs like *görünmez adam* [ɟørynmæzadam] 'invisible man' vs. *görünme zamanı* [ɟørynmezamanɯ] 'time to appear'. This quasi-contrastivity between [e] and [æ] may have contributed to reinforce the lexicalization of exceptions to the lowering process.

Another morphological conditioning on lowering could be analogical in nature. The harmonically front form of the aorist suffix allomorph /-Ar/ (occurring after monosyllabic roots ending in a consonant) is *-er*, with regular lowering: [ær]. Lowering in *-mez* might thus have been modelled on the phonological shape of the other aorist morpheme. In turn, *pekmez* is one of the very few Turkish roots that end with the syllable *-mez*, which happens to be identical in form to the

negative aorist suffix.[2]

2.4. DERIVED WORDS. In at least one instance, the application of /e/ lowering exhibits variability within a morphologically derived environment. One of the very few prefixes of Turkish is an adjective (and occasionally adverb) intensifier that consists in a copy of the first consonant (if present) and first vowel of the root, followed by [p] if the root is vowel initial, and either from the set {p, s, m, r} (Göksel & Kerslake 2005; Kornfilt 1997) if the root is consonant initial. Which consonant is selected is subject to fairly complex factors (see Tang & Akkuş (2023) for a recent analysis of them.)

followed by an additional consonant to a degree, based on the consonants of the root (Göksel & Kerslake 2005: 90): *sarı* 'yellow' / *sapsarı* 'bright yellow', *temiz* 'clean' / *tertemiz* 'clean as a pin'. When this morpheme precedes the adjective *pembe* 'pink', some speakers produce the form *pespembe* with [æ] in both syllables: [pæspæmbe]. While the second [æ], occurring before /m/, is predicted by the lowering rule, the first [æ] occurs before /s/, unexpectedly for a purely allophonic rule triggered by a sonorant coda — for those speakers, the requirement of vowel identity between prefix and root appears to treat [æ] as a phoneme distinct from [e].

2.5. EFFECT OF WORD FREQUENCY. With regard to the effect of word frequency, to our knowledge the role it plays in the process in question has never been investigated. Since more frequent words appear to be more likely to undergo processes leading to lexical split than low-frequency words do (Labov 1994), evidence of a correlation between the amount of vowel lowering and the lexical frequency of the word it occurs in would support the hypothesis that we are dealing with an instance of 'lexical' phonological split. Therefore, our last research goal is investigating the relation, if any, between vowel lowering and lexical frequency.

2.6. LEXICAL EXCEPTIONS. Some exceptions to the lowering rule are completely unmotivated phonologically, such as [e] in *hem* 'both' (all other instances of /e/ before a coda /m/ are lowered).

3. Methodology. In order to clarify the phonetic quality of the surface realizations of Turkish /e/, we created a list consisting of 107 items (105 words in isolation, plus the phrases *görün-mez adam* and *görünme zamanı* mentioned above) containing one or more instances of this vowel. Since our goal was to investigate the realization of /e/ in each environment, our items include words where /e/ occurs in open and closed syllables; among closed syllables, in some of them the coda is a sonorant, in others the coda is an obstruent. Finally, to compare lowering at word boundaries with word-internal lowering, we included items where /e/ occurs in the final syllable and items where /e/ is in a non-final syllable (Table 1).

The list was presented to nine Turkish-speaking participants (4 male, 5 female, average age 22), who were asked to read each word aloud as it appeared on a computer screen. The stimuli were presented in isolation in a randomized order, and each stimulus occurred three times in total. The subjects were free to move to the next item by pressing a button on the keyboard. The recordings were made in a quiet room using an AT2020 USB external cardioid condenser microphone connected to a fanless laptop. The audio files contained 4200[3] tokens, which were manually annotated in Praat (Boersma & Weenink 2023) marking the onset and offset of each /e/

[2] We would like to thank an anonymous reviewer for suggesting the potential impact of analogy on lowering in *pekmez* and *-mez*.

[3] This number is higher than the product of 107 by 9 by 3 because several items have more than one /e/.

token at zero-crossing and labeled. Afterwards, a Praat script (Lennes n.d.) extracted F1 and F2 values at midpoint, and these values were Lobanov-normalized[4] (Lobanov 1971). The resulting data were then analyzed and visualized using R (R Core Team 2021), with the `dplyr` package for data manipulation (Wickham et al. 2023) and `ggplot2` to create our plots (Wickham 2016).

The role of word frequency was investigated using the BOUN Corpus (Sak et al. 2008), which consists of four sub-corpora. Three of these sub-corpora are derived from leading Turkish newspapers, while the fourth is a general collection of Turkish web pages. Its creators refer to the combined data from these sub-corpora as the BOUN Corpus. The total corpus contains 423 million tokens, 4.1 million word forms, and 48637 roots. Sak et al. (2008) applied a morphological parser to analyze the words in the corpus. For this paper, the frequency results provided by this parser were used, totaling around 383 million tokens.

Some of the words we used for our experiment have zero frequency in the BOUN corpus (that is, they never occur there). In order to deal with this problem and use the information provided by these items, we applied a Laplace transformation to our frequency data. It consists in adding 1 to each frequency count and increasing the total corpus size by the number of word types observed. See Brysbaert & Diependaele (2013) for a discussion of its appropriateness and superiority over other strategies to deal with zeroes in corpora.

Syllable Type	Coda	Environment	Examples
Closed	Sonorant	Internal / Final	*erkek*, *tünel*
Closed	Obstruent	Internal / Final	*asker*, *tüfek*
Open	Ø	Internal / Final	*gebe*, *tepe*
Closed	Obstruent	Final	*gel-mez*, *görün-mez*
Closed	Sonorant	Lexical Exceptions	*elbise*, *pekmez*
Open / Closed	Obstruent / Ø	Word Boundary	*görün-mez adam*, *görünme zamanı*

Table 1. The environments targeted

4. Results and Discussion. In this section, we present and discuss our measurements in relation to the model of lexical split introduced in section 2, and we fit a statistical model to our data.

4.1. HEIGHT AND DISCRETENESS OF THE OUTPUT OF THE LOWERING RULE. We measured F1 values — the most important acoustic correlate of vowel height — by syllable type for the vowels measured. Our data are in line with previous claims (Göksel & Kerslake 2005; Gopal & Nichols 2022) that when /e/ occurs before a sonorant coda, it is significantly lowered, essentially being realized as [æ]. The average F1 frequency for female speakers is 807 Hz, while for male speakers it is 640 Hz. These values are broadly comparable to those reported for [æ] in other languages. For example, Hillenbrand et al. (1995) found average F1 frequencies of 669 Hz for female speakers and 588 Hz for male speakers of American English. Similarly, Mokari & Werner (2016) reported averages of 788 Hz for female speakers and 605 Hz for male speakers of Azerbaijani. Iivonen & Harnud (2005) documented an F1 value of approximately 630 Hz for a male Finnish speaker. Furthermore, as Figure 1 shows, the lowered allophone is near-categorically distinct from [e], which occurs before obstruent codas and in open syllables; F1 values in syllables

[4] Lobanov's (1971) original formula used the root mean square deviation; following common practice, we used the standard deviation.

closed by a sonorant only marginally overlap with those in the other two environments (higher (normalized) F1 value = lower vowel, lower (normalized) F1 value = higher vowel; the negative aorist *-mez* and the word *pekmez* are not included in this figure; the bimodal distribution in open syllables is due to the different effect of word-final vs. non-word-final position — see Figure 5; the bimodal distribution in syllables closed by a sonorant is due to inter-speaker variability in the lowering of words like *ben* 'I,' *sen* 'you,' and *en* 'most' — see section 1).

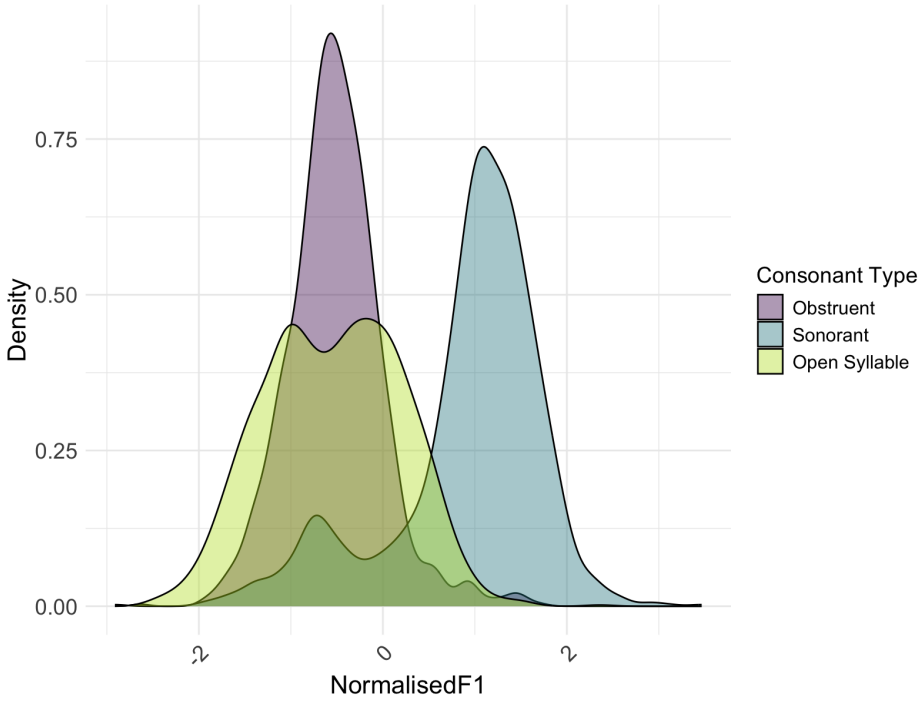

Figure 1. Effects of syllable and consonant type on F1

4.2. NATURE OF THE CONDITIONING ENVIRONMENT. We also wanted to investigate if more fine-grained consonantal effects exist beyond the broad sonorant-obstruent dichotomy. Figure 2 demonstrates that, sonority aside, the specific quality of the coda consonant has a very limited impact on the amount of lowering. /e/ has very similar F1 values before any coda obstruent (considering Turkish /j/ not to be a sonorant phonologically, as we have argued in fn. 1), whereas F1 is significantly and uniformly higher before all coda sonorants (or in the final syllable of the word *pekmez* and in the suffix *-mez*).

4.3. LOWERING BEFORE /z/. Our measurements also show that for all our nine speakers the vowel /e/ in the negative aorist *-mez* and in *pekmez* is roughly as low as it is before a sonorant coda, whereas in any other word /z/ codas are not correlated with lower F1 values of /e/ (Figure 3; the sonorant outliers with low F1 values are the lexical exceptions that are mentioned in section 1).

Figure 4 is specifically devoted to the over-application of /e/ lowering in some syllables ending in /z/. Our results confirm Gopal & Nichols's (2022) report of lowering in *pekmez* and the negative aorist *-mez*. However, in our data we did not find any significant lowering in the word *merkez*, contrary to their findings. All the other words in our dataset having an /ez/ rhyme do not

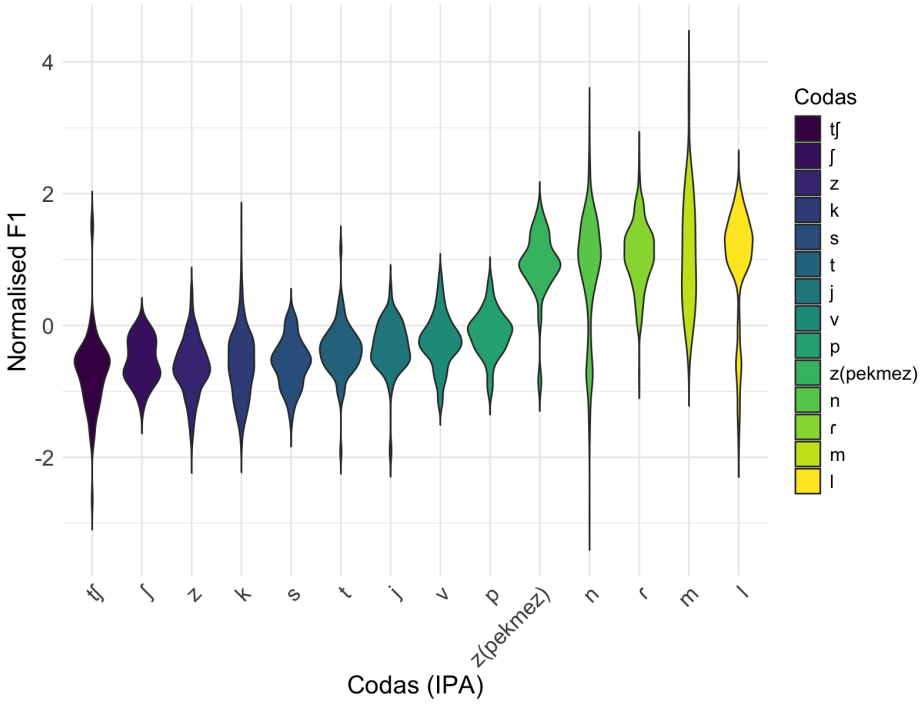

Figure 2. F1 values of the mid-front unrounded vowel before different coda consonants

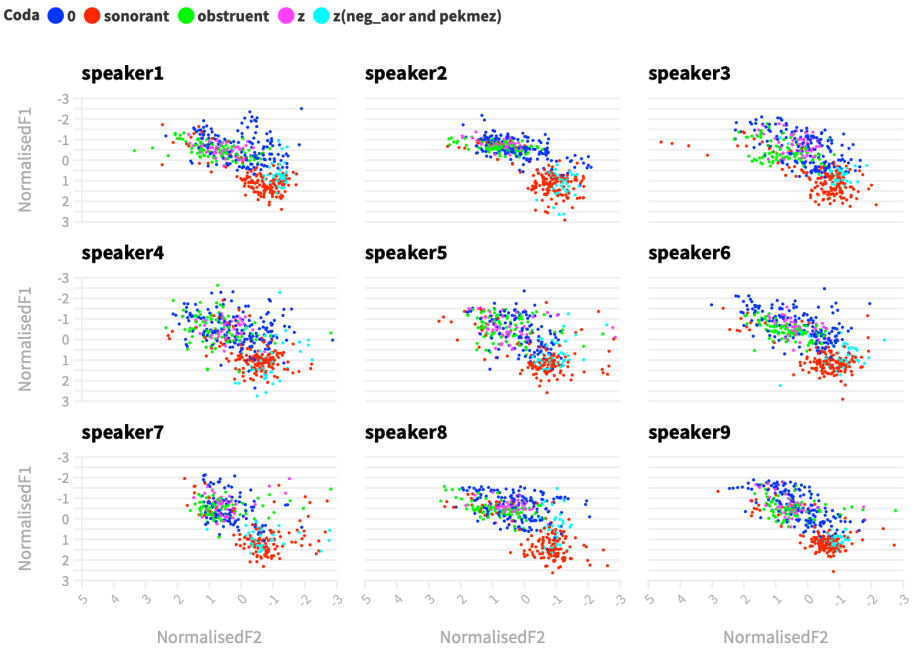

Figure 3. /e/ tokens of all speakers

show lowering either.

4.4. LOWERING IN WORD-FINAL OPEN SYLLABLES. With regard to syllables not closed by a sonorant consonant, our results support Göksel & Kerslake's (2005) claim that /e/ is lowered

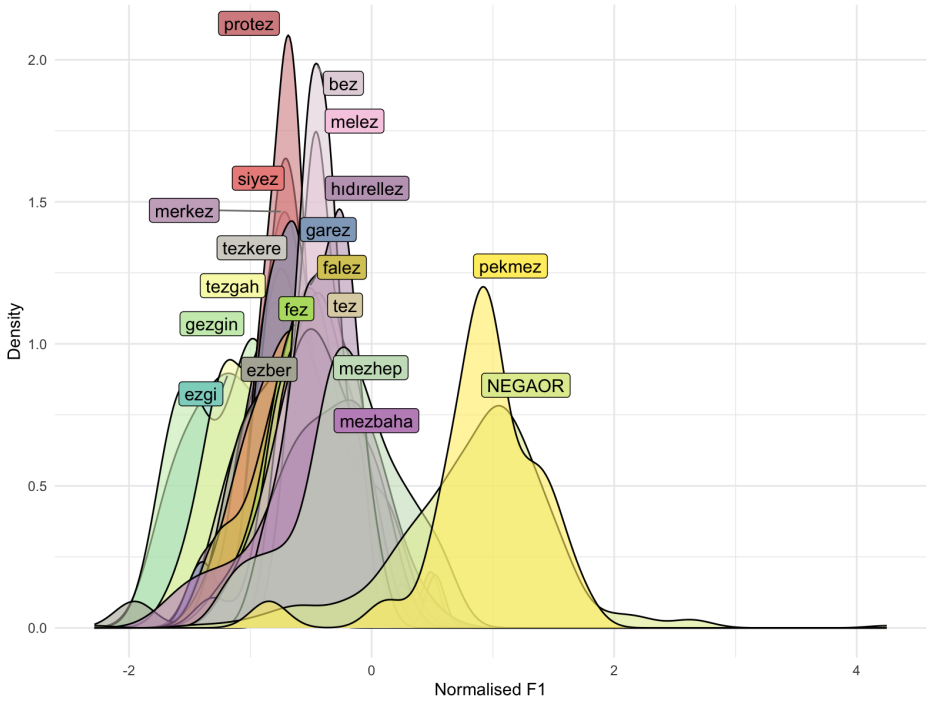

Figure 4. Lowering before /z/

to [ɛ] in word-final open syllables (Figure 5). Compared to lowering before sonorant codas, the output of this process is not as lowered and is much more gradient, and in all likeness is a part of a more general process of vowel lowering in word-final position (according to Göksel & Kerslake 2005, all vowels except /a/ and /o/ are lowered). As for word-final closed syllables ending in an obstruent, their vowel is slightly higher than in word-final open syllables, but slightly lower than in non-final closed syllables ending in an obstruent.

4.5. EFFECT OF WORD FREQUENCY. Our data show a positive correlation between the amount of lowering /e/ and the log-normalized corpus frequency of the morpheme the vowel belongs to: higher frequency words tend to have a lower vowel (Figure 6).

4.6. STATISTICAL MODELING. A linear mixed-effects model was fitted to the data using the `lmer` function in the R `lme4` package (Bates et al. 2015). The type of coda consonant following /e/ (or /e/ being within *pekmez* or the negative aorist, in the case of lexically specific lowering), syllable position within the word, and lexical frequency were treated as fixed effects, while speaker and word were treated as random effects. We treated each consonantal category as a level of the coda consonant factor, rather than using a broader categorization such as Obstruent/Sonorant, to examine whether individual consonant qualities influence lowering. Grouping all obstruents and all sonorants together could obscure the impact of finer-grained consonantal properties, should they be relevant. The reference level for this factor is [ʧ], i.e. the environment where /e/ has the highest realization in our data. Since we hypothesize that only lowering to [æ] is sensitive to lexical frequency, and only lowering to [ɛ] is restricted to word-final syllables, we also included the interaction between syllable position and the category following the vowel (which includes open syllables among its levels), and the interaction between the cate-

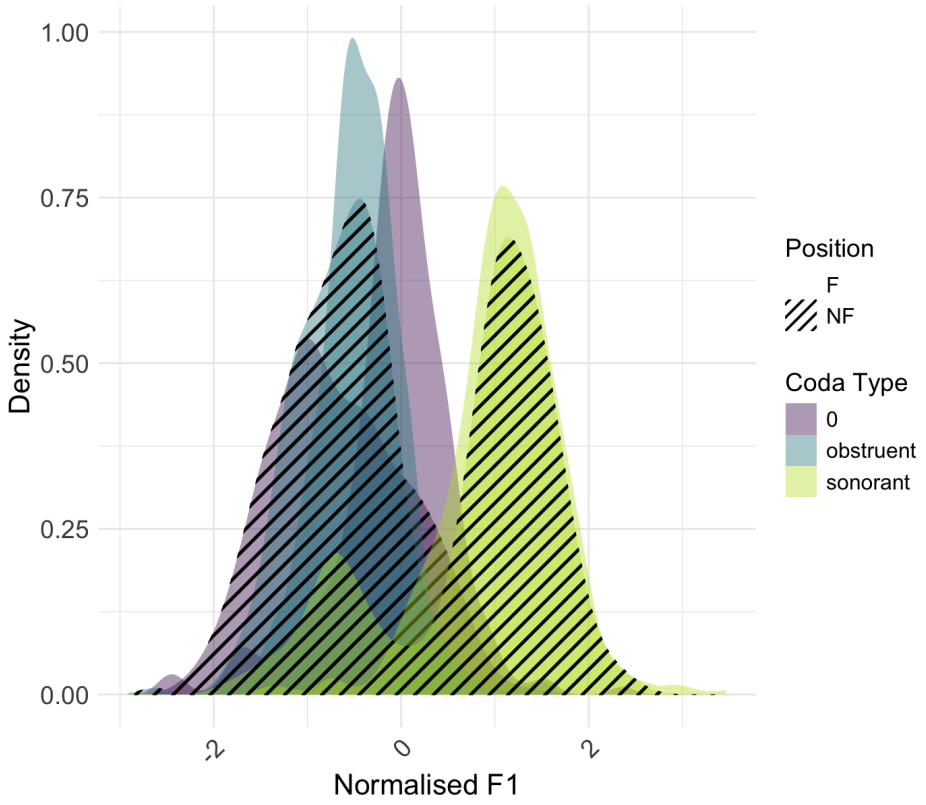

Figure 5. Height of /e/ in every environment excluding *pekmez* and *-mez*

gory following the vowel and lexical frequency. The `drop1()` function was used to assess the significance of each term in the model by performing single-term deletions and calculating the corresponding changes in model fit. Based on its results, the interaction between the category following the vowel and lexical frequency was removed, while all the other variables were retained.

The resulting model (Table 2) confirms that the F1 value of the non-high front unrounded vowel is significantly positively correlated with the following consonant being a sonorant, or the vowel appearing in the word *pekmez*, or in the suffix *-mez*. Word frequency also plays a significant role, with an increase in lexical frequency of a word predicting a relatively higher F1 value in its non-high front unrounded vowel(s) (i.e., a lower realization). Lastly, the interaction between syllable position within a word and the category of the following segment is generally not significant. The minor negative impact of /l/ in non-final position on vowel height arises because two exceptional words that do not exhibit lowering — *elbise* and *Belçika* — contain this consonant.

5. Conclusions and directions for further research. Our findings confirm previous claims that Turkish /e/ is systematically lowered to [æ] before sonorant codas, forming in this environment a set of phonetic realizations that is fundamentally discontinuous from those occurring in open syllables or closed syllables ending in an obstruent.

Our data also confirm that lowering unexpectedly occurs in *-mez* 'NEG.AOR' and the final syllable of *pekmez* 'molasses'. Unlike in the data discussed by Gopal & Nichols (2022), we did not find evidence of lowering in the final syllable of *merkez* 'center'. However, this is consistent with the already observed across-speaker variability of exceptional under-application of lowering;

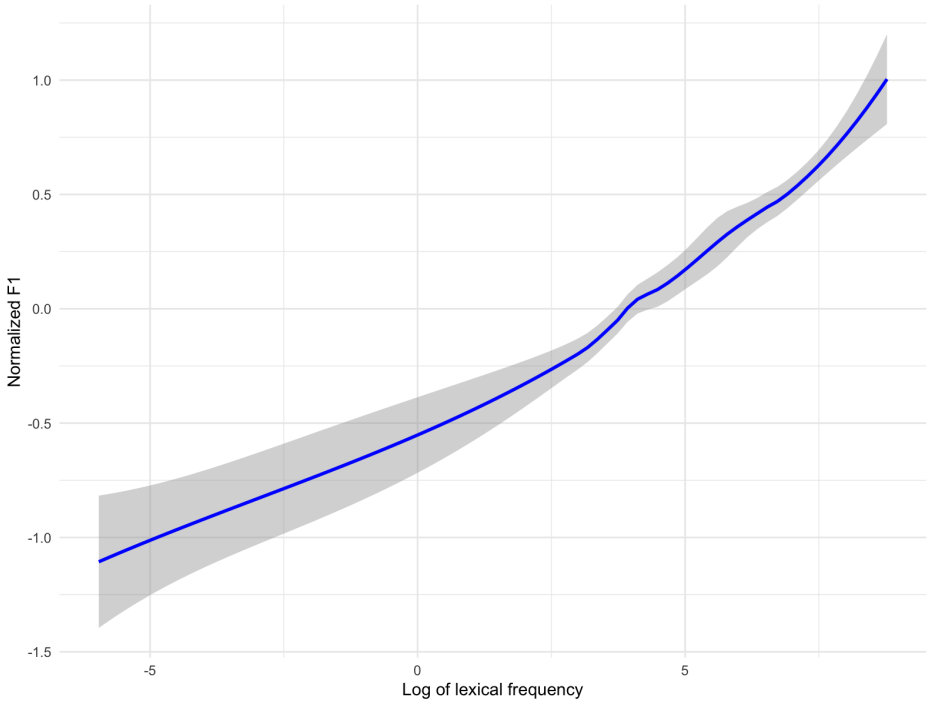

Figure 6. Correlation between lexical frequency and F1

not every speaker exhibits all the observed lexical exceptions.

More generally, instances of under- and over-application suggest that the process is influenced by a complex set of phonetic, lexical, and morphological factors. In particular, we identified one previously unreported aspect of /e/ lowering in Turkish, i.e., a positive correlation between the lexical frequency of a word and the amount of lowering. This property aligns with a well-documented tendency in linguistic change. High-frequency words often lead 'lexical' phonological splits, which may also be affected by non-phonological boundaries (Labov 1994). Thus, these findings support the hypothesis of an in-progress phonologization of the allophone [æ], which seems to be acquiring a marginal contrastive status. It may be useful to add that in Gopal & Nichols's (2022) longitudinal data the distance between the allophones of /e/ appears to be increasing in younger generations, which is consistent with the idea that we are observing an ongoing diachronic change.

As for the diachronic mechanism originating this embryonic split of the Turkish phoneme /e/, the most plausible explanation is in terms of lexical diffusion (in fact, according to Labov 1994 'lexical' phonological splits can only arise through lexical diffusion). We hypothesize an initially purely allophonic lowering rule of /e/ before coda sonorants. Contact with Turkish dialects in which [e] contrasts with a lower front unrounded vowel (see section 1) likely introduced lexical exceptions to the allophonic rule, as [e] can occur before coda sonorants in these dialects. The possibility of exceptions to the rule may in turn have triggered the word-by-word reanalysis of some occurrences of [æ] as a separate phoneme, when favoured by analogy, lexical frequency and/or presence of word boundaries (see section 2).[5]

[5] It is worth noting that occurrences of [æ] before [z] are unlikely to represent a diachronic preservation. The vowel of the negative aorist *-mez* derives from the East Old Turkic mid front unrounded vowel *ä*, like the majority of in-

	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	-1.1370	0.5094	70.08	-2.2320	0.0288	*
Following Cons k	-0.0103	0.5121	71.52	-0.0200	0.9841	
Following Cons z	0.3588	0.5286	76.85	0.6790	0.4993	
Following Cons s	0.4439	0.5327	81.77	0.8330	0.4071	
Following Cons ʃ	0.5923	0.7962	54.40	0.7440	0.4602	
Following Cons j	0.6343	0.6549	57.05	0.9690	0.3368	
Following Cons p	0.7902	0.5283	80.92	1.4960	0.1386	
Following Cons v	0.9048	0.6213	62.42	1.4560	0.1503	
Following Cons t	1.0020	0.6260	63.56	1.6010	0.1143	
Following Cons 0 (open syll)	1.4740	0.5105	71.22	2.8880	0.0051	**
Following Cons n	1.6340	0.5146	71.90	3.1750	0.0022	**
Following Cons m	1.6820	0.7840	53.33	2.1460	0.0365	*
Following Cons l	1.8330	0.5397	73.21	3.3960	0.0011	**
Following Cons r	2.2100	0.5176	73.75	4.2690	0.00006	***
pekmez	1.8330	0.5456	85.41	3.3590	0.0012	**
negative aorist	1.9010	0.7830	52.43	2.4280	0.0187	*
Position in Word NonFinal	0.0770	0.4300	845.00	0.1790	0.8579	
Log of lexical frequency	0.0001	0.0001	64.08	2.1270	0.0373	*
Following Cons k: Word NonFin	-0.0188	0.4698	751.20	-0.0400	0.9680	
Following Cons z: Word NonFin	0.2301	0.4717	756.30	0.4880	0.6258	
Following Cons s: Word NonFin	-0.3318	0.4791	1035.00	-0.6930	0.4888	
Following Cons 0 (open syll): Word NonFin	-0.6607	0.4344	871.30	-1.5210	0.1287	
Following Cons n: Word NonFin	-0.0313	0.4534	892.90	-0.0690	0.9450	
Following Cons m: Word NonFin	1.1140	0.7673	89.44	1.4510	0.1502	
Following Cons l: Word NonFin	-0.8851	0.4999	488.90	-1.7710	0.0772	.
Following Cons r: Word NonFinal	0.1768	0.4652	783.00	0.3800	0.7039	

Table 2. Statistical model

Finally, it is worth emphasizing that our study is only preliminary. It has only touched upon some aspects of this lowering process, leaving many others unexplored. In particular, the role of morphologically derived environments, sociolinguistic variables affecting lowering, and the acceptability of lexical exceptions remain to be understood.

References

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1). https://doi.org/10.18637/jss.v067.i01.

stances of /e/ in modern Turkish (see Johanson 2021 for a sketch of the phonological system of East Old Turkic). Instances of modern Turkish /e/ with a different historical origin stem not from a lower vowel but a higher one — a 'closed e' that was likely still present in Ottoman Turkish (Kerslake 2021: 177-178). This vowel, still attested in many parts of Anatolia (see section 1), has merged with the reflex of *ä* in modern Turkish (Csató & Johanson 2021: 195). As for the words *merkez* and *pekmez*, the former is certainly a borrowing from Arabic, while the etymology of the latter is less certain.

Boersma, Paul & David Weenink. 2023. *Praat: Doing phonetics by computer*. Praat software. http://www.praat.org/. Version 6.3.18.

Brendemoen, Bert. 2021. Turkish Dialects. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic Languages*, 224–230. Routledge.

Brysbaert, Marc & Kristof Diependaele. 2013. Dealing with zero word frequencies: A review of the existing rules of thumb and a suggestion for an evidence-based choice. *Behavior Research Methods* 45(2). 422–430. https://doi.org/10.3758/s13428-012-0270-5.

Canalis, Stefano, Semra Özdemir, Utku Türk & Ümit Tunçer. 2024. The phonological nature of the Turkish front glide. In Fatma Hülya Özcan Önder, Tuncay Karalık, Bayram Çibik, İlknur Civan Biçer & Samet Deniz (eds.), *Selected essays on Turkish Linguistics: The Anadolu Meeting. Proceedings of the 20th International Conference on Turkish Linguistics*, 61–78. Wiesbaden: Harrassowitz Verlag.

Clements, George N. & Engin Sezer. 1982. Vowel and Consonant Disharmony in Turkish. In Harry van der Hulst & Norval Smith (eds.), *The Structure of Phonological Representations Part 2*, 213–256. De Gruyter. https://doi.org/10.1515/9783112423325-007.

Csató, Éva Á. & Lars Johanson. 2021. Turkish. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic Languages*, 195–223. Routledge.

Erguvanlı Taylan, Eser. 2015. *The Phonology and Morphology of Turkish*. Istanbul: Boğaziçi University Press.

Göksel, Aslı & Celia Kerslake. 2005. *Turkish: A Comprehensive Grammar*. New York: Routledge.

Gopal, Deepthi & Stephen Nichols. 2022. Apparent-time change in Turkish front-mid-vowel allophony. Talk at LAGB 2022, 15 September.

Hillenbrand, James, Laura A. Getty, Michael J. Clark & Kimberlee Wheeler. 1995. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97(5). 3099–3111.

Iivonen, Antti & Huhe Harnud. 2005. Acoustical comparison of the monophthong systems in Finnish, Mongolian and Udmurt. *Journal of the International Phonetic Association* 35(1). 59–71.

Johanson, Lars. 2021. East Old Turkic, 132–144. Routledge.

Kerslake, Celia. 2021. Ottoman Turkish. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic Languages*, 174–194. Routledge.

Kiparsky, Paul. 2015. Phonologization. In Patrick Honeybone & Joseph Salmons (eds.), *Handbook of Historical Phonology*, 563–580. Oxford: Oxford University Press.

Kornfilt, Jaklin. 1997. *Turkish Grammar*. London: Routledge.

Labov, William. 1994. *Principles of Linguistic Change: Internal Factors*. Oxford: Blackwell.

Lennes, Mietta. n.d. Formant-logger. Praat script. Available: https://github.com/lennes/spect/blob/master/scripts/collect_formant_data_from_files.praat.

Lobanov, B. M. 1971. Classification of Russian Vowels Spoken by Different Speakers. *The Journal of the Acoustical Society of America* 49(2B). 606–608. https://doi.org/10.1121/1.1912396.

Mokari, Payam Ghaffarvand & Stefan Werner. 2016. An acoustic description of spectral and temporal characteristics of Azerbaijani vowels. *Pozna Studies in Contemporary Linguistics* 52(3). 503–518.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing Vienna, Austria. https://www.R-project.org/.

Sak, Haşim, Tunga Güngör & Murat Saraçlar. 2008. Turkish language resources: Morphological parser, morphological disambiguator and web corpus. In *Advances in natural language processing*, 417–427. Springer Berlin Heidelberg.

Tang, K. & F. Akkuş. 2023. Identity Avoidance in Turkish Partial Reduplication: Feature Specificity and Locality. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 14(1). 1–52. https://doi.org/10.16995/labphon.6459.

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Romain Franois, Lionel Henry, Kirill Mller & Davis Vaughan. 2023. *dplyr: A grammar of Data Manipulation*. https://dplyr.tidyverse.org. R package version 1.1.4, https://github.com/tidyverse/dplyr.