

High vowel shortening in Turkish

Elif E. Sarmış & Stefano Canalis*

Abstract. High vowel shortening is a common phenomenon that usually depends on purely phonetic factors. This study aims to present the findings of an experiment we conducted on 5 native speakers. The experiment was designed to ascertain whether Turkish shortens high vowels and if so, to understand whether shortening is controlled by phonological or phonetic factors. Our data show little correlation between a vowel's F_1 and its duration within the high ([+high], /i, u, y, ʊ/) and non-high ([-high], /a, e, o, ø/) categories in all environments (stressed and unstressed, in open and closed syllables, in all consonantal environments). On the other hand, a comparison of the duration of [+high] and [-high] vowels reveals that non-high vowels are rarely as short as, or shorter than, high vowels. These results suggest that vowel shortening in Turkish is not a purely phonetic phenomenon; what matters for this process to apply is the phonological category of the vowel. All [+high] vowels are shortened, largely irrespective of variation in their F_1 ; conversely, [-high] vowels are not shortened, largely irrespective of variation in their F_1 .

Keywords. Turkish phonology, vowel shortening, vowel weakening, high vowels

1. Introduction. High vowels have long been reported to be more susceptible to phonetic weakening/reduction than non-high vowels. This encompasses a number of different but similar phenomena, notably including shortening, devoicing, and in more extreme cases, complete deletion of high vowels. This paper presents some preliminary findings of an ongoing research on high vowel weakening in Turkish. The following sections will be almost exclusively confined to shortening, but high vowel shortening and devoicing are often closely intertwined (see for example the discussion about Japanese below); while the present study will not discuss devoicing in detail, it will be included in the final report (Sarmış 2022), and we briefly outline some general properties of both processes in this section.

All other things being equal, it has been observed in many languages that higher vowels tend to be shorter and have a greater tendency to devoice than lower vowels. Such phenomena are well documented for languages such as Northern Greek (Topintzi & Baltazani 2012), Japanese (Jun & Beckman 1993; Tanner et al. 2019), Korean (Jun & Beckman 1994), and Uzbek (Zhou & Vogel 2018). Indeed, both shortening and devoicing of high vowels are so common that they are considered to be phonetic universals (cf. for example Maddieson 1999: 622). Their source, however, is a matter of debate: do they occur simply due to articulatory or aerodynamic factors or are they phonological?

* We would like to thank two reviewers for their careful reading of our manuscript and their comments and remarks, which helped us improve the article. We would also like to thank the participants of TU+6 for their helpful comments, in particular Barış Kabak, John Kingston and Jonathan Washington, as well as Noyan Dokudan and Pavel Logačev for their invaluable help in the design of our experimental materials, and Furkan Atmaca, Elena Guerzoni, Muhammed İleri, and Aref Milani for their precious comments and suggestions. The usual disclaimers apply. This project was funded by the *Boğaziçi University Bilimsel Araştırma Projeleri Koordinatörlüğü* [Pr. No: 16021, ‘Syllabic phonology’].

Authors: Elif E. Sarmış, Boğaziçi University (elifsarmis@gmail.com) & Stefano Canalis, Boğaziçi University (stefano.canalis@boun.edu.tr).

In most languages, both processes appear to result from inherent articulatory properties of speech. With regard to devoicing, the main cause seems to be aerodynamic. Vowels are generally voiced segments, therefore requiring vocal fold vibration to be produced. For vocal fold vibration to occur, air must pass through the glottis, which is only possible if the subglottal pressure is higher than the supraglottal pressure. The increased constriction in the production of higher vowels causes oral airflow to be more impeded, raising the supraglottal pressure. This makes it harder to achieve the trans-glottal pressure differential necessary to start or maintain voicing (Maddieson 1999: 622). Higher vowels, therefore, devoice more frequently than lower vowels.

This may be especially noticeable next to voiceless consonants. The absence of vocal fold vibration in the preceding and/or following consonant, combined with the difficulty in starting voicing in a high vowel and the susceptibility of the latter to shortening, may lead to a complete or nearly complete absence of vocal fold vibration during the vocalic gesture for a high vowel.

High vowel shortening (hereafter HVS), on the other hand, appears to be the result of physiological/articulatory factors. Lower vowels require more lowering of the tongue and jaw, thus also requiring more time than a higher vowel to reach their articulatory target. As a corollary, lower vowels are longer in duration than higher vowels, making higher vowels shorter than their lower counterparts – what has been known as ‘intrinsic vowel duration’ at least since Lehiste (1970: 18) (see also, among many others, Catford 1977: 196–7, Maddieson 1999: 622).

Therefore, HVS and devoicing typically appear to be purely phonetic (mechanical, intrinsic, articulatorily or aerodynamically based, and gradient). However, it has become increasingly clear that drawing the exact boundary between phonetics and phonology is anything but simple. Some phonological processes previously thought to be fully categorical have been argued to be at least in part gradual. To name but one, some American English speakers categorically assimilate alveolar nasals as in e.g., *ban cuts* to the place of the following velar consonant, but other speakers only partially assimilate, gradiently retaining aspects of the alveolar closure (Ellis & Hardcastle 2002). Conversely, facts that seemed to depend only on physiological, acoustic, or aerodynamic constraints are at least in part governed by phonology in some languages: they are learned, controlled, and (semi)categorical. For example, vowels have been found to be longer before voiced than before voiceless consonants in almost all the languages in which the interaction between vowel duration and consonantal voicing has been examined, suggesting that some universal phonetic factor underlies this asymmetry (although its precise nature is debated – see, e.g., Chen 1970). At the same time, some cases of vowel lengthening before voiced consonants show language-specific properties, presumably controlled by the phonology of the languages in question. For instance, in English the magnitude of the effect is larger than in most other languages (Chen 1970: 138). Furthermore, in American English the mean duration of vowels preceding flaps is longer if the underlying consonant is /d/ rather than /t/, and vowel duration does not significantly change when the flap is voiceless (as in 19% of all flaps in Fox & Terbeek’s 1977 data). In other words, while “[t]he durational differentials between vowels before /d/-flaps as opposed to /t/-flaps are on the order of magnitude of the *universal* difference [...] argued to be due to [...] physiological constraints [...], this kind of explanation simply cannot hold for the present data, since the vowel duration differences are caused by abstract phonological distinctions” (Fox & Terbeek 1977: 33; emphasis in the original).

HVS and devoicing are no different in this respect. While plausibly they are exclusively phonetic in many, perhaps most languages, it has been argued that in some languages they are at least partially phonologized. Thus, Tauberer & Evanini (2009) report that dialectal variation in

first formant frequency across North American English dialects appears to be independent of intrinsic vowel duration, which therefore seems to be a target specified in the grammar rather than the result of universal physiological constraints. Toivonen et al. (2015) found that, while higher vowels are generally shorter than lower vowels in English and Swedish, tokens of the same vowel differing in the value of F_1 do not display the same tendency, which is unexpected if the effect were purely mechanical. They conclude that each English and Swedish vowel has a phonologized duration target.

A similar state of affairs holds true for high vowel devoicing. While in most languages it seems to depend purely on aerodynamic causes, enhanced by gestural overlap when the high vowel is adjacent to voiceless consonants, in some languages there seems to be a phonological rule devoicing high vowels. The difficulty in drawing a boundary between phonetic and phonological vowel reduction is aptly exemplified by the debate about Japanese devoiced high vowels between voiceless consonants, with some authors seeing them as the gradient by-product of the gestural overlapping of the laryngeal gestures (Jun & Beckman 1993), and others interpreting them as the output of a phonologically driven categorical rule (Tsuchida 1997). Interestingly, Japanese high vowel devoicing and shortening seem to interact as two facets of a more general weakening rule. Tanner et al. (2019: 1) report that “[Japanese] high vowels are substantially shorter before voiceless consonants, whilst non-high vowels do not exhibit comparable shortening”, which leads them to argue that in Japanese duration is phonologically controlled to favor devoicing.

Even though modern standard Turkish (referred simply as ‘Turkish’ henceforth) is one of the most studied languages within the Turkic family, the phonology and phonetics of Turkish are still relatively understudied areas of research, and high vowel weakening (hereafter HVW) is not an exception. The only experimental works entirely devoted to Turkish HVW are Jannedy (1994, 1995) and Ay & Bekâr Uzun (2015), and they are mainly focused on devoicing (the latter only discussing /u/). To our knowledge, HVS has never been investigated comprehensively; a few works (e.g., Şayli 2002, Şayli et al. 2002, Şayli & Arslan 2003, Arısoy et al. 2004, Aydınlı 2018, Ergenç & Bekâr Uzun 2020: 107–10, Lorenzen 2021) report measurements of Turkish vowel durations, and do consistently find that high vowels are shorter than non-high ones (see for example Table 1 below, which reproduces measurements from Arısoy et al. 2004), but do not extensively comment on this finding.

	/a/	/e/	/o/	/ø/	/u/	/i/	/u/	/y/
Mean	74.0	76.0	84.2	82.4	57.4	49.6	53.0	58.8
Std. dev.	15.8	13.2	14.0	20.2	20.6	17.7	14.0	13.4

Table 1: Means and standard deviations of vowel durations in word-initial syllables (in ms), adapted from Arısoy et al. 2004: 6.

The aim of this paper is thus twofold: first, we want to expand the empirical data on Turkish HVS. Second, we want to understand whether it is purely phonetic or if it is controlled by a (semi)categorical phonological process. We conducted an experiment with five native speakers of Turkish, and our experimental observations suggest that HVS is an allophonic rule controlled by phonology and a manifestation of a more general HVW process.

While the results of our experiment suggest that Turkish has both shortening and devoicing in high vowels, due to space limitations this paper will be mainly focused on HVS.

2. Phonological vs. phonetic height. In the following sections, we will try to argue that HVS has been phonologized in Turkish; shortening is no longer the consequence of physiological, mechanical constraints, but the output of a phonological rule that shortens [+high] vowels.

Such a hypothesis of course implies that it is possible to tell phonetic shortening apart from phonological shortening. Probably all languages have – all else equal – shorter high(er) vowels, but if shortening were controlled by phonology its properties should at least partially differ from purely phonetic shortening. We suggest four criteria that may be useful to separate phonetic from phonological shortening.

A first clue may be the size of shortening. While HVS is plausibly universal, the reported amount of durational difference is fairly small in languages that have only phonetic shortening, usually below or close to the threshold of perceptibility (although see Lehiste 1970: 18). We would expect that if shortening were phonological the durational difference between high and non-high vowels would be higher.

A second criterion may be the categorical vs. gradient nature of shortening. Tongue and jaw position are articulatory continua; therefore, vowel duration should vary more or less continuously with their position if only the former two influenced the latter. Conversely, if shortening targeted phonological categories, we would expect more categorical durational differences between vowels. For instance, only /a/ would be longer than any other vowel if [-low] vowels were shortened, while mid and low vowels would have similar durations if [+high] vowels were shortened.

Third, articulatory aspects that affect vowel tongue height should have an impact on duration in languages where shortening is phonetic, but they would plausibly be irrelevant (or at least less relevant) if shortening were phonological. For example, the hard palate is higher than the soft palate, which means that the highest point reached by the body of the tongue is typically lower for [u] (and more generally back vowels) than for [i] (and more generally front vowels of the same height category as a back one), to the extent that, for instance, not only American English /i/ but also /ɪ/ have a higher tongue position than /u/ (Ladefoged & Maddieson 1996: 283–284). Given the slightly lower tongue position of back vowels when compared to front vowels having the same phonological height, it would be expected that they take a slightly longer time than front vowels to be produced if intrinsic vowel duration solely depended on the time required to reach their articulatory target.

Finally, another way to separate phonetic and phonological height is provided by the intrinsically variable phonetic instantiation of phonological categories (see Toivonen et al. 2015 for a similar argument with regards to the correlation between vowel height and duration in English and Swedish). The phonetic realization of (the height and duration of) a vowel is never exactly the same; even ignoring major allophonic differences, it may be influenced, to name only a few parameters, by the surrounding consonants, stress, intonational pattern, word length, the dialect and idiolect of the speaker, and so on. Even two repetitions of the same vowel in the same environment uttered by the same speaker will not be completely identical. All these variables can affect the actual position of the tongue. If HVS were phonetic in Turkish, the duration would be influenced by the actual position of the tongue, not by a vowel's phonological category ([+high] or [-high]). To put it differently, the correlation between vowel duration and height would be expected to hold within as well as between vowel categories. For instance, lower-than-usual phonetic realizations of /e/ should be longer than higher-than-usual tokens of the same vowel, and the latter should be only moderately longer (if not as long as) lower-than-usual realizations of /i/.

Conversely, if the phonological category of vowels were the main determinant of their duration, variation in the actual tongue height among /e/ tokens should be largely unrelated to their variation in duration; higher-than-usual tokens of /e/ would be on average as long as lower-than-usual tokens and would not be durationally closer to tokens of /i/. All else equal, vowel duration would depend on the vowel being [+high] or [-high], [+high] vowels being shortened and [-high] ones being not, independently of their actual tongue height. This process would be relatively categorical, Turkish high vowels being fairly regularly shorter than non-high vowels regardless of their actual tongue position.

The vowel system of Turkish shows a few such discrepancies between phonetic and phonological height. Its inventory consists of the vowels /i, y, u, e, ø, o, a/. Phonologically, it seems sufficient to only distinguish between high (or [+high]) and non-high (or [-high]) vowels to account for both the vowel contrasts and the vowel alternations of Turkish. Regarding contrasts, the phonetically low /a/ differs from any other vowel in at least one of the values of [back], [round], and [high], making its specification as [+low] redundant. As for phonological processes, height is relevant in a few phonological processes in Turkish, but none of them refers to the difference between [+low] and [-low] vowels; only high vowels undergo rounding harmony, epenthetic vowels are high (see for example Clements & Sezer 1982), and high vowels optionally assimilate to non-high ones in hiatuses (Kabak 2007). Again, the only phonologically relevant height property of all vowels seems to be being either [+high] or [-high], and indeed /a/ is the [+back] counterpart of /e/ in palatal harmony (for example, the two allomorphs of the plural suffix are *-lar* and *-ler*).

There is thus a lot of variation in tongue position and F_1 that is not determined by the [+high] or [-high] specification of a vowel in Turkish. Just to mention some of them, /a/ is significantly lower than any mid vowel. /e/ has three major allophones differing in height: [ɛ] occurs word-finally, [æ] occurs in closed syllables where the coda is a sonorant consonant, [e̞] elsewhere (Göksel & Kerslake 2005: 10). Turkish /u/, traditionally described as a high back unrounded vowel, seems to actually be slightly lower (as well as more central) than the other (truly) high vowels /i/, /u/ and /y/ (Kılıç & Ögüt 2004). Hence, there is only a rather coarse correspondence between the [\pm high] specification of Turkish vowels and the actual tongue and jaw position, which provides empirical material to test which of them influences duration the most.

(1) **If HVS were phonetic in Turkish:**

- Vowel duration depends on the amount of time required to move articulatory organs
 - lower vowels need more time to reach their target, making them longer
- What matters is the actual position of the tongue, not the vowel's phonological category

If HVS were phonological in Turkish:

- Everything else equal, vowel duration depends on the vowel being [+high] or [-high]
 - /i, u, y, u/ as a group are shortened while /a, e, o, ø/ are not, irrespective of their actual tongue height
- What matters is the phonological category of the vowel, not the actual position of the tongue

3. Methods. We conducted an experiment to explore HVS in Turkish, testing high (/i, y, u, u/) and non-high (/e, ø, a, o/) vowels in open and closed syllables, while controlling the preceding and following environment and stress position.

Participants. The experiment was conducted on 5 Turkish native speakers (four female, one male) aged between 20 and 55 years (the experiment was actually conducted on 18 speakers in

total [12 F, 6 M] but we are still in the process of analyzing the data). They were all born and raised in Turkey and speak modern standard Turkish ('Istanbul Turkish').

Materials. Our experimental items were minimal (or near-minimal, when the exact word did not exist) pairs of Turkish words differing in the height of one of their vowels (e.g., *kısa* 'short' vs. *kaşa* 'safe (N)', *bilge* 'sage' vs. *belge* 'document').

In addition to vowel height, we also manipulated other parameters that cross-linguistically commonly affect vowel duration. The consonants preceding and following the vowel included voiceless and voiced consonant, as well as the 'zero environment' (word initial vowel, Ø, e.g., *inik* 'deflated' vs. *enik* 'puppy, kitten') to test the effects of voicing. Our experimental items included both open and closed syllables to measure the effect of syllable type on HVS (e.g., *deri* 'skin' vs. *derdi* 'trouble.POSS').

We also added stress as a condition. In Turkish, primary stress falls on the final syllable in most words (however, there are some lexical exceptions, e.g., *Bébek* 'a district in Istanbul' versus *bebék* 'baby'). Nevertheless, some morphemes shift stress position, one example being the negation marker *-mA*, which causes stress to fall on the syllable before it. We used the negation marker to manipulate stress position in the experiment, creating minimal or near-minimal pairs that only differed in their stress pattern (e.g., *karmá* 'mixed' versus *kárma* 'don't mix'). All items had two to four syllables and were relatively high-frequency words. To minimize the impact of utterance-level prominence on vowel duration, items were elicited within the carrier sentence shown in (2) below.

- (2) *Oku-duğ-um cümle-de _____ kelime-si var-dı / yok-tu*
 read-NMLZ-1SG sentence-LOC _____ word-CM exist-PAST / exist.NEG-PAST
 "The sentence I read contained / did not contain the word _____."
 (NMLZ: nominalizer, CM: compound marker)

Data collection. Participants were seated in a silent room and were presented with a self-paced reading task where they were asked to read a sentence that contained the experimental item and then to complete the carrier sentence out loud, repeating this five times. The data were recorded with the software Audacity (Audacity Team 2021) using an external microphone (Audio-Technica AT2035) connected to a laptop through a preamplifier (PreSonus AudioBox USB 96).

Data analysis. While participants were asked to repeat the carrier sentence five times, only four iterations for each participant were analyzed, discarding the first utterance. The data were analyzed using Praat (Boersma & Weenink 2021). We used a vowel's first formant (F_1) as a proxy for its phonetic height, as it is universally accepted that it is the acoustic correlate of vowel height (see, e.g., Ladefoged & Johnson 2011: 220–1, Ladefoged & Maddieson 1996: 284). Vowel height and F_1 are inversely correlated: the higher a vowel, the lower is its F_1 .

We used two Praat scripts (Crosswhite 2003, Welby 2003) with minimal changes to extract the formant and duration values for each vowel after manually segmenting the audio files. The formant script extracts F_1 , F_2 , and F_3 values from the midpoint of each sound and writes the values in a text file. We hand-checked items in Praat when the formant values provided by the scripts were unlikely for a given vowel. A total of 98 items were used in this experiment and 1633 tokens were analyzed from 5 participants (362 tokens from speaker A, 340 from B, 319 from C, 292 from D, and 320 from E).¹

¹ The varying number of tokens for each participant is due to some tokens being discarded because of background noise and malfunctioning equipment.

When vowels were devoiced their formant structure was faint and irregular, and when they were deleted there was no periodicity at all. Since our analysis depended on the formant values of the vowels, devoiced and deleted segments were discarded from the dataset. In total, we discarded 122 tokens, leaving us with 1511 tokens (1 /e/, 27 /i/, 21 /u/, 25 /y/ and 48 /ʉ/ tokens were discarded; incidentally, these numbers show that while devoicing/deletion of high vowels is not uncommon – 121/878, or 13.8 percent of them were devoiced/deleted – it is practically untested in non-high vowels).

4. Results. Before discussing the influence of (phonetic or phonological) height on vowel duration, we briefly review here the other main parameters that affect vowel duration in Turkish, since their effect must be factored out from duration to understand the role of height alone.

First, all else equal, Turkish vowels are shorter before as well as after voiceless consonants (Table 2). Second, longer words tend to have shorter vowels (Table 3)².

	before			after	
	voiced	voiceless	no cons.	voiced	voiceless
Average V duration	54	44	64	53	49

Table 2: Average duration of the vowels (in ms) in different voicing environments

	2 syllables	3 syllables	4 syllables
Average V duration	60	49	33

Table 3: Average duration of the vowels (in ms) depending on word length

Both effects are unsurprising. As mentioned above, the lengthening effect of consonant voicing on vowel duration seems to be universal or near-universal (Chen 1970). The compression effect of word length on duration is also well documented (Lehiste 1970: 40).

A third factor affecting the duration of Turkish vowels is the syllable structure: vowels are longer in closed than in open syllables (Table 4). Unlike the previous two findings, this is quite uncommon cross-linguistically; if anything, in many languages, vowels are shortened in closed syllables (Maddieson 1999: 628). However, our results are in line with those of previous studies on vowel duration in Turkish (e.g., Lahiri & Hankamer 1988, Jannedy 1995: 26–29), which also report longer vowels in closed syllables.

Finally, we found practically no durational difference between stressed and unstressed vowels (Table 5). This result is typologically unexpected as well, as duration is one of the most common and most reliable correlates of stress – in Matthew Gordon and Timo Roettger’s recent cross-linguistic survey it distinguishes stress in 65 of the 72 languages in their database (Gordon & Roettger 2017: 5). However, also in this case our data agree with the findings of previous studies. For example, Levi (2005) and Vogel et al. (2016: 34) report negligible differences in the duration of stressed vs. unstressed Turkish vowels.

	open	closed
Average V duration	36	52

Table 4: Average duration of the vowels (in ms) in different syllable types

² The seemingly large negative effect of word length on vowel duration is due to the low number of four-syllable words in our word list, most of them having high vowels. The statistical models we present in the next section show that the impact of word length on duration is relatively minor.

	stressed	unstressed
Average V duration	51	50

Table 5: Average duration of the vowels (in ms) in stressed and unstressed positions

Moving to the main subject of our research, the table and the bar plot below show the average duration of each Turkish vowel. It is apparent from our data that Turkish displays HVS. Vowel height affects vowel duration in Turkish, high vowels being substantially shorter than non-high vowels (Table 6 and Figure 1); this is true of each of our subjects (Table 7). What remains to be determined is the phonetic or phonological nature of this influence.

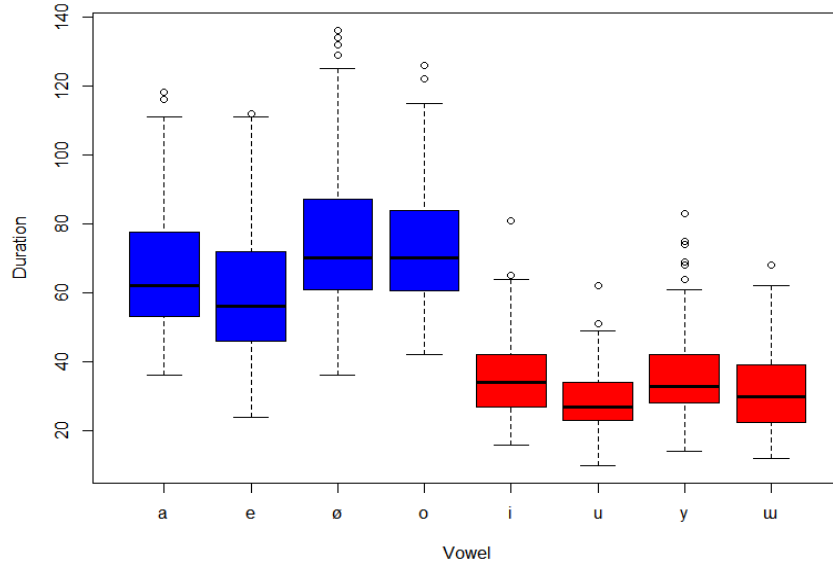


Figure 1: The duration of individual vowels (mean of all speakers)

	/a/	/e/	/o/	/ø/	/u/	/i/	/y/	/ɯ/
Avg. V duration	67	59	74	75	32	36	29	36

Table 6: Average duration of individual vowels (in ms, mean of all speakers)

	Speaker A	Speaker B	Speaker C	Speaker D	Speaker E
High	32	39	32	32	27
Mid	80	83	63	58	63
Low	76	79	58	63	58

Table 7: Average duration of vowels (in ms) for each speaker

5. Analysis and discussion. We mentioned some possible criteria to distinguish between phonetic and phonological shortening in the preceding section. HVS in Turkish can usefully be compared to vowel shortening in Italian, a language that, to our knowledge, has never been claimed to have phonological HVS. The vowel inventory of Standard Italian is /a, ε, ɔ, e, o, i, u/. Data from Esposito (2002) show that its higher vowels are shorter than lower vowels (Table 8).

	/i/	/u/	/e/	/o/	/ε/	/ɔ/	/a/
Mean	112	117	137	141	149	157	150
Std. dev.	20	24	27	29	27	31	28

Table 8: Average duration of Italian vowels (in ms). Data from Esposito (2002: 204)

The first criterion we discussed is the magnitude of the difference in duration. The average duration of all Italian high vowels combined is 114.5 ms, while that of non-high vowels is 146.8, that is 28.2% longer than high vowels. By contrast, the average duration of all Turkish high vowels combined is 34 ms, while that of non-high vowels is 73 ms, that is 114.7% longer.³

The second criterion we discussed is the gradualness of the correlation between height and duration. Italian high vowels are shorter than mid-high vowels, which in turn are shorter than mid-low vowels. Only /a/ marginally departs from this correlation, being slightly shorter than /ɔ/ (but significantly longer than any high or mid-high vowel, and as long as or slightly longer than /ɛ/). By contrast, the transition is more categorical in Turkish; the difference between high vs. mid vowels is very pronounced, whereas all non-high vowels have similar durations, even if their phonetic height ranges from mid to low.

The third criterion we proposed is whether vowel duration is correlated with other articulatory dimensions that affect tongue height. As mentioned in section 2, it is well known that back vowels typically have a lower tongue position than front vowels having a similar F_1 value; given this slightly lower tongue position of back vowels, it would be expected that they would take a slightly longer time than front vowels to be produced if intrinsic vowel duration solely depended on the time required to reach their articulatory target. Indeed, Italian back vowels are moderately but consistently longer than front vowels of the same height, which may be explained by the articulatory account of HVS. Interestingly, the same does not hold in Turkish, which shows no clear correlation between vowel backness and duration.

All these facts are suggestive of a qualitative difference between Turkish and Italian HVS, but none of them proves decisively that Turkish HVS is determined by the phonology of the language (for instance, there is no obvious threshold at which durational differences can be assumed to become phonological).

We believe that our fourth criterion – the variable phonetic instantiation of one and the same phonological category – may help separate more clearly the phonetic and the phonological impact of vowel height on duration in Turkish. We present below statistical models that examine how the two alternative hypotheses fit our data.

In Tables 9 and 10 we compare two linear mixed-effects models.⁴ One of them includes as a predictor phonetic-only information about the height of a vowel (that is, its F_1 ⁵), whereas the other includes phonological-only information about vowel height (that is, whether a vowel is [+high] or [-high]). Both models also include other fixed effects that plausibly affect vowel duration, i.e., voicing of the preceding consonant, voicing of the following consonant, open or closed syllable type, number of syllables of the word the vowel is in, and stress. ‘Speaker’ is treated as a random effect.

³ The large difference in absolute duration between Italian and Turkish is largely ascribable to stress. Esposito (2002) measured stressed open syllables, which are lengthened in Italian. On the contrary, our data combine stressed and unstressed vowels, and even more importantly, Turkish stress seems to have little or no impact on vowel duration.

⁴ Modeling was conducted using the *lme4* (Bates et al. 2015) R package. There is no obvious and universally agreed-upon way to estimate p -values in linear mixed models, given the uncertainty in calculating degrees of freedom. We adopt Satterthwaite’s approximation as implemented in the function *summary* in the R package *lmerTest* (Kuznetsova et al. 2017).

⁵ In order to minimize variation among speakers caused by purely physiological differences, their vowels’ F_1 values were normalized adopting Lobanov’s transformation (Lobanov 1971), as implemented in the function *norm.lobanov* in the *vowels* R package (Kendall & Thomas 2010).

Coefficients	Estimate	Std. Error	df	t value	p-value
Intercept	66.30	3.11	5.11	21.34	< 0.001 ***
Prev seg: voiceless	-13.67	1.00	1481.40	-13.66	< 0.001 ***
Prev seg: zero	1.81	1.49	1481.20	1.22	0.22
Fol seg: voiceless	-4.80	0.98	1481.20	-4.88	< 0.001 ***
Syl count: 3	-10.13	1.04	1481.70	-9.70	< 0.001 ***
Syl count: 4	-17.13	1.53	1481.34	-11.22	< 0.001 ***
Syl type: open	-5.87	1.70	1481.20	-3.45	< 0.001 ***
Stress: yes	0.64	0.89	1481.60	0.72	0.47
Normalized F_1	11.74	0.47	1481.05	24.91	< 0.001 ***

Table 9: Model one, phonetic height

Coefficients	Estimate	Std. Error	df	t value	p-value
Intercept	43.42	3.60	4.74	12.49	< 0.001 ***
Prev seg: voiceless	-6.93	0.82	1481.10	-8.42	< 0.001 ***
Prev seg: zero	3.96	1.19	1481.09	3.32	< 0.001 ***
Fol seg: voiceless	-5.70	0.79	1481.08	-7.24	< 0.001 ***
Syl count: 3	-6.09	0.85	1481.22	-7.21	< 0.001 ***
Syl count: 4	-8.28	1.26	1481.03	6.58	< 0.001 ***
Syl type: open	-12.18	1.37	1481.07	-8.89	< 0.001 ***
Stress: yes	1.70	0.71	1481.28	2.40	< 0.05 *
Vowel height: [-high]	33.26	0.78	1481.25	42.46	< 0.001 ***

Table 10: Model two, phonological height

In both models, voicing of the preceding and following consonants is positively correlated with longer vowel duration, while the effect of the absence of a consonant before the vowel is small. A lower number of syllables also increases duration, while stress has no significant effect in the first model and a minor and marginally significant one in the second. With regard to vowel height, the two models show that both phonetic and phonological height are correlated with duration. Irrespective of which hypothesis is more correct, this result is unexpected, since phonetic and phonological height are different but intertwined entities: phonologically high vowels have on average lower F_1 than [-high] vowels, and vowels with lower F_1 tend to be phonologically [+high] vowels. What is more interesting for our hypothesis is the amount of variance in vowel duration that is accounted for when either dimension of vowel height is included in the model while all the other variables are kept constant. If vowel duration were mainly affected by phonetic height, the first model would provide a better fit to the durational data. On the contrary, if phonological height controlled vowel duration in Turkish, the latter model would provide a better fit. We estimated the R^2 of both models and of each fixed predictor within them (which approach to adopt to obtain R^2 from mixed models is still debated; we used the *r2beta* function in the R package *r2glmm*, Jaeger 2017). As the tables below show, using a categorical specification of vowel height (Table 12) instead of raw phonetic height (that is, F_1 in Table 11), provides a better fit to the data; the R^2 value of phonological height is 0.549, while that of (normalized) F_1 in the other model is 0.295. The R^2 value of the whole model using phonological rather than the phonetic height is higher as well: 0.677 vs. 0.495.

Effect	R^2	Upper Conf. Lim.	Lower Conf. Lim.
Model	0.495	0.527	0.465
Normalized F_1	0.295	0.331	0.260
Prev seg voi	0.136	0.169	0.106
Syl count	0.092	0.121	0.066
Fol seg voi	0.016	0.031	0.006
Syllable type	0.008	0.019	0.002
Stress	0.000	0.005	0.000

Table 11: R^2 of the model in (9)

Effect	R^2	Upper Conf. Lim.	Lower Conf. Lim.
Model	0.677	0.699	0.655
Phonol height	0.549	0.578	0.520
Prev seg voi	0.074	0.102	0.052
Syl type	0.051	0.074	0.031
Syl count	0.041	0.063	0.024
Fol seg voi	0.034	0.054	0.018
Stress	0.004	0.013	0.000

Table 12: R^2 of the model in (10)

A further way to estimate the relative impact of phonetic and phonological height on duration is provided by data where phonological height remains constant, but phonetic height varies. They would reveal the contribution of phonetic height alone to duration. The model in Table 10 does show a statistical correlation between the F_1 value of a vowel and its duration, but it includes vowels having different [\pm high] values. However, if Turkish vowel shortening mainly depended on the amount of time required to move articulatory organs, with lower vowels requiring more time to reach their articulatory target, this should hold true not only when comparing [+high] to [-high] vowels, but also when comparing vowels within the same phonological category, as each category includes tokens with markedly different F_1 values. Turkish non-high vowels range in phonetic height from /e, o, ø/ (usually described as phonetically mid, so [e̞, o̞, ø̞] in a narrower transcription) to phonetically low /a/, and /e/ has mid-low as well as near-low allophones, providing a wide range of F_1 variation within the category of [-high] vowels. The phonetic tokens of [+high] vowels do not always have the same phonetic height either; as mentioned above, Turkish /u/ seems to be slightly lower than /i, y, u/. Finally, all vowels, even when they have a very similar mean duration and F_1 , display a fairly large amount of variability along these parameters.

Therefore, the articulatory account of HVS would predict the correlation between F_1 and duration to be equally strong within as across the [-high] and [+high] vowel categories. Tables 13 and 14 show two mixed linear models investigating this; in the former, F_1 is the predictor of the duration of [+high] vowels, while the latter does the same for [-high] vowels (these two models also include the same fixed effects as above except for stress, and ‘speaker’ as a random effect).

They show that the correlation between F_1 and duration is quite weak within the [-high] and [+high] vowel categories; there is no statistically significant correlation between F_1 and duration within [+high] vowels, while the correlation within [-high] vowels is statistically significant but much weaker than the one found in Table 9 (the size of the effect in Table 9 is less than a fifth of that in Table 10). Hence, once phonetic height variability is separated from phonological vowel height, the contribution of the former to duration turns out to be quite modest.

Coefficients	Estimate	Std. Error	df	t value	p-value
Intercept	43.93	2.50	6.69	17.59	< 0.001 ***
Prev seg: voiceless	-4.97	0.92	740.34	-5.37	< 0.001 ***
Prev seg: zero	-1.06	1.59	740.19	-0.66	0.51
Fol seg: voiceless	0.07	0.92	740.55	0.07	0.94
Syl count: 3	-10.22	0.95	740.32	-10.73	< 0.001 ***
Syl count: 4	-11.15	1.27	740.10	-8.76	< 0.001 ***
Syl type: Open	-4.01	1.26	740.35	-3.20	< 0.01 **
Normalized F_1	1.18	0.80	740.37	1.48	0.14

Table 13: Phonetic height within [-high] vowels

Coefficients	Estimate	Std. Error	df	t value	p-value
Intercept	77.34	5.07	4.34	15.24	< 0.001 ***
Prev seg: voiceless	-6.26	1.48	730.02	-4.24	< 0.001 ***
Prev seg: zero	3.78	1.68	730.05	2.25	< 0.05 *
Fol seg: voiceless	-10.40	1.32	730.05	-7.87	< 0.001 ***
Syl count: 3	-3.17	1.25	730.08	-2.54	< 0.05 *
Syl count: 4	-6.15	4.08	730.09	-1.51	0.13
Syl type: Open	-20.10	3.35	730.08	-6.01	< 0.001 ***
Normalized F_1	2.21	0.65	730.27	3.39	< 0.001 ***

Table 14: Phonetic height within [+high] vowels

6. Conclusion. This paper aimed to present the findings of an experimental study conducted on five native Turkish speakers. We explored the shortening of Turkish high vowels in various environments to figure out whether this shortening is caused by purely articulatory and aerodynamic factors or is instead governed by phonology. Although the current study was based on data from only five speakers, the results appear to support the conclusion that Turkish high vowels are shortened. This is anything but surprising since HVS is supposedly universal. We have also seen in the previous sections that several facets of our experimental evidence converge to support an additional, less obvious finding: Turkish HVS is only in a minimal part the mechanical product of articulatory constraints. On the whole, vowel duration is not directly affected by the degree of jaw opening; whatever correlation between these two properties may be observed, the [\pm high] feature specification of a vowel reveals a much more cogent correlation with duration.

The magnitude of the effect as well as its near-categorical output are more compatible with a phonological process targeting [+high] vowels than with a gradient, purely articulatory mechanism. The [\pm high] feature specification of a vowel also accounts for more variance in vowel duration than its actual F_1 value, and additionally, F_1 shows a weak correlation with vowel duration within the two phonological categories [+high] and [-high].

We thus feel it legitimate to suppose that the phonology of Turkish includes an optional allophonic rule shortening [+high] vowels. In other words, the speaker has a shorter durational target when a vowel is [+high]. We say the rule is optional and allophonic because a [+high] vowel as long as a [-high] one is uncommon/not prototypical rather than utterly ungrammatical (for example, in our data the longest /i/ token by speaker A is 59 ms long, while her shortest /a/ token is 54 ms long). However, the not entirely categorical nature of HVS can be best be accounted for in terms of phonological optionality, rather than as the outcome of a purely articulatory, gradient pressure. We would also like to point out that other aspects of Turkish phonology suggest that high vowels are ‘weaker’ than non-high ones, thus providing converging evidence for our claim

that high vowels are phonologically reduced.⁶ First, phonologically [+high] vowels are more easily prone to assimilation than non-high vowels; they assimilate to a preceding non-high vowel in hiatuses (Kabak 2007), and only high vowels undergo rounding harmony. Furthermore, Turkish epenthetic vowels are high, and epenthetic vowels are typically the most unmarked vowels in an inventory (Rice 2007).

Finally, we would like to mention some directions for future research. We are currently collecting data from more speakers, and we plan to include high vowel devoicing – that we observed in our data – in our account of Turkish high vowel weakening. The optional nature of the process raises the question whether the use of an experimental setting might partially suppress it, and consequently whether recordings of spontaneous speech would reveal even more reduction, possibly being subject to sociolinguistic factors. We therefore intend to also examine spontaneous speech corpora of spoken Turkish.

References

- Arısoy, Ebru, Levent M. Arslan, Mine N. Demiralp, Hazım K. Ekenel, Meltem Kelepir, H. Mesut Meral, A. Sumru Özsoy, Ömer Şaylı, Oytun Türk & Beste Can Yolcu. 2004. Duration of Turkish vowels revisited. *12th International Conference on Turkish Linguistics (ICTL 2004)*, Dokuz Eylül University, Izmir, Turkey.
- Audacity Team. 2021. *Audacity* [Computer program]. Version 3.0.2, retrieved from <https://www.audacityteam.org>.
- Ay, Sıla & İpek Pınar Bekâr Uzun. 2015. A devoicing analysis of vowel [i̥] in voiceless consonant surroundings. In Deniz Zeyrek, Çiğdem Sağın Şimşek, Ufuk Ataş & Joehen Rehbein (eds), *Ankara Papers in Turkish and Turkic Linguistics*. 2–11. Wiesbaden: Harrassowitz Verlag. <https://doi.org/10.2307/j.ctvc770nr>.
- Aydınlı, Fatma Esen, Maviş Emel Kulak Kayıkçı & Nilda Suslu. 2018. Temporal and frequency characteristics of Turkish vowels in laryngectomized speakers: Preliminary study. *Medeniyet Med J*, 34. 149–159. <https://doi.org/10.5222/MMJ.2019.42744>.
- Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1). 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Boersma, Paul & David Weenink. 2021. *Praat: doing phonetics by computer* [Computer program]. Version 6.1.16, retrieved from <http://www.praat.org/>.
- Catford, John C. 1977. *Fundamental problems in Phonetics*. Bloomington: Indiana University Press.
- Chen, Matthew. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica*, 22. 129–159. <https://doi.org/10.1159/000259312>.
- 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 II)*. 213–255. Dordrecht: Foris.
- Crosswhite, Katherine. 2003. *Duration logger* [Praat script]. Retrieved from http://phonetics.linguistics.ucla.edu/facilities/acoustic/duration_logger.txt

⁶ We would like to thank Barış Kabak for suggesting us to explore the connection between HVS and other phonological processes of Turkish.

- Ellis, Lucy & William J. Hardcastle. 2002. Categorical and gradient properties of assimilation in alveolar to velar sequences: Evidence from EPG and EMA data. *Journal of Phonetics*, 30. 373–396.
- Ergenç, İclal & İpek Pınar Bekâr Uzun. 2020. *Türkçenin ses dizgesi (2nd edn)*. 107-110. Ankara: Seçkin Press.
- Esposito, Anna. 2002. Vowel height and consonantal voicing effects. *Phonetica*, 59. 197–231.
- Fox, Robert A. & Dale Terbeek. 1977. Dental flaps, vowel duration and rule ordering in American English. *Journal of Phonetics*, 5(1). 27–34.
- Gordon, Matthew & Timo Roettger. 2017. Methodological issues in the study of word stress correlates. *Linguistics Vanguard* 3(1).
- Göksel, Aslı & Celia Kerslake. 2005. *Turkish: A comprehensive grammar*. London: Routledge.
- Jaeger, Byron. 2017. *r2glmm: Computes R Squared for mixed (multilevel) models* [R package]. Version 0.1.2, software resource: <https://CRAN.R-project.org/package=r2glmm>.
- Jannedy, Stefanie. 1994. Prosodic and segmental influences on high vowel devoicing in Turkish. [MA Thesis, Ohio State University]. OSU Library Repository.
- Jannedy, Stefanie. 1995. Gestural phasing as an explanation for vowel devoicing in Turkish. *OSU Working Papers in Linguistics*, 45.
- Jun, Sun-Ah & Mary Beckman. 1993. A gestural-overlap analysis of vowel devoicing in Japanese and Korean. Paper presented at the 67th Annual Meeting of the Linguistic Society of America, Los Angeles.
- Jun, Sun-Ah & Mary Beckman. 1994. Distribution of devoiced high vowels in Korean. *Proceedings of the 1994 International Conference on Spoken Language Processing*, 1. 479–482. Yokohama: Acoustic Society of Japan.
- Kabak, Barış. 2007. Hiatus resolution in Turkish: An underspecification account. *Lingua*, 117(8). 1378–1411. <https://doi.org/10.1016/j.lingua.2006.04.010>.
- Kendall, Tyler & Erik R. Thomas. 2010. *Vowels: Vowel manipulation, normalization, and plotting in R* [R package]. Version 1.1, software resource: <https://CRAN.R-project.org/package=vowels>.
- Kılıç, Mehmet Akif & Fatih Ögüt. 2004. A high unrounded vowel in Turkish: Is it a central or back vowel? *Speech Communication*, 43. <https://doi.org/10.1016/j.specom.2004.03.001>.
- Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen. 2017. lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13). 1–26. <https://doi.org/10.18637/jss.v082.i13>.
- Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Cambridge, MA: Blackwell.
- Ladefoged, Peter & Keith Johnson. 2011. *A course in Phonetics (6th edn)*. Wadsworth: Cengage Learning.
- Lahiri, Aditi & Jorge Hankamer. 1988. The timing of geminate consonants. *Journal of Phonetics*, 16(3). 327–338.
- Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.
- Levi, Susannah V. 2005. Acoustic correlates of lexical accent in Turkish. *Journal of the International Phonetic Association*, 35(1). 73–97. <https://doi.org/10.1017/S0025100305001921>.
- Lobanov, Boris M. 1971. Classification of Russian vowels spoken by different listeners. *Journal of the Acoustical Society of America*, 49.

- Lorenzen, Janne. 2021. Full and partial vowel devoicing in Turkish and the role of prosodic prominence. Paper presented at the annual meeting of the 4th Phonetics and Phonology in Europe (PaPE2021), Barcelona (Online).
- Maddieson, Ian. 1999. Phonetic universals. In William J. Hardcastle & John Laver (eds.), *The handbook of Phonetic sciences*. 619–639. Oxford, UK: Blackwell.
- Rice, Keren. 2007. Markedness in phonology. In Paul de Lacy (ed.), *The Cambridge handbook of Phonology*. 79–97. Cambridge: Cambridge University Press.
- Sarmış, Elif E. (2022). High vowel weakening in Turkish. [MA Thesis, Boğaziçi University].
- Şayli, Ömer. 2002. Duration analysis and modelling for Turkish text-to-speech synthesis. [MS Thesis, Boğaziçi University].
- Şayli, Ömer, Levent M. Arslan & A. Sumru Özsoy 2002. Duration properties of the Turkish phonemes. *11th International Conference on Turkish Linguistics*. Famagusta, Northern Cyprus.
- Şayli, Ömer & Levent M. Arslan. 2003. Türkçe'deki seslerin süre özellikleri. *Dilbilim Araştırmaları*. 15–26, Istanbul, Turkey: Boğaziçi University Press.
- Tanner, James, Morgan Sonderegger & Francisco Torreira. 2019. Durational evidence that Tokyo Japanese vowel devoicing is not gradient reduction. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.00821>.
- Tauberer, Joshua & Keelan Evanini. 2009. Intrinsic vowel duration and the post-vocalic voicing effect: Some evidence from dialects of North American English. *Proceedings of the 10th Annual Conference of the International Speech Communication Association*.
- Toivonen, Ida, Lev Blumenfeld, Andrea Gormley, Leah Hoiting, John Logan, Nalini Ramlakhan & Adam Stone. 2015. Vowel height and duration. *Proceedings of the 32nd West Coast Conference on Formal Linguistics*. 64–71. Somerville: Cascadilla Proceedings Project.
- Topintzi, Nina & Mary Baltazani. 2012. The acoustics of high-vowel loss in a Northern Greek dialect and typological implications. In Philip Hoole, Lasse Bombien, Marianne Pouplier, Christine Mooshammer & Barbara Kühnert (eds.), *Consonant clusters and structural complexity*. 369–393. Berlin: De Gruyter Mouton.
- Tsuchida, Ayako. 1997. Phonetics and phonology of Japanese vowel devoicing. [Ph.D. dissertation, Cornell University].
- Vogel, Irene, Angeliki Athanasopoulou & Nadia Pinkus. 2016. Prominence, contrast, and the functional load hypothesis: An acoustic investigation. In Jeffrey Heinz, Rob Goedemans & Harry van der Hulst (eds.), *Dimensions of Phonological stress*. 123–167. Cambridge: Cambridge University Press.
- Welby, Pauline. 2003. *Get formant* [Praat script]. Retrieved from <https://www.uvic.ca/humanities/linguistics/assets/docs/praat/get-formant.praat>
- Zhou, Chen & Irene Vogel. 2018. High vowel reduction/deletion in Uzbek. *Proceedings of the 3rd Workshop on Turkish, Turkic, and the languages of Turkey (TU+3)*. Los Angeles: UCLA.