2022. Proceedings of the Workshop on Turkic and Languages in Contact with Turkic 7. 164 - 171.

Continuant Devoicing in Turkish: Reanalysis of Turkish Word-Final Devoicing

Ezgi Sanıyar & Ümit Can Tunçer*

Abstract. Previous research shows that Turkish voiced stops and affricates undergo devoicing in word-final position. There are some grammars claiming that rhotic [ɾ] is also affected by this phenomenon (Göksel & Kerslake, 2005 and Taylan, 2015) as well as liquid sonorants ([l] and [ʎ]) which can optionally undergo word-final devoicing, at least in the Istanbul dialect of Turkish (Kornfilt, 1990). This study tests the continuant consonants in Turkish for their devoicing properties in different environments. These environments include preceding phonemes of different voicing qualities, word-final position, sentence-final position, and isolation. Our experiment examines all the voiced continuant sounds of Turkish that can naturally occur in word-final positions ([ɾ], [l], [ɫ], [ʎ], [j], [v], [β], [z], and [ʒ]). The experiment involves 10 Turkish-speaking participants from different regional backgrounds in Turkey. Sets of words and sentences focusing on the target phones and environments were presented to the participants to read out loud while recording. The findings suggest that word-final devoicing does extend to the [+cont] consonants of Turkish in the word-final position, though it shows variance among the participants. This phonological process seems to be optional unlike stop devoicing in Turkish which is a systematic process. Another finding, also observed by Nichols (2016) for the rhotic only, was that the majority of the instances where "word-final" devoicing occurred, was in fact exclusively in utterance-final position. Environments that can be considered word-final but not utterance final, yielded mostly fully voiced consonants, showing evidence for [+cont] consonants having their own separate set of phonological rules and environments for devoicing than stop consonants. Overall, this analysis suggests an environment of devoicing for [+cont] consonants (voiced fricatives and non-nasal sonorants) in Turkish and reanalyzes the rhotic [ɾ] devoicing which was previously described as word-final only, similar to the stop consonants.

Keywords. phonetic analysis; Turkish; continuant devoicing; sonorants; fricatives; word-final

1. Introduction. Turkish has a well-documented case of word-final devoicing of voiced stops ([-cont], [-son]) (Hulst&Weijer, 1991, Kopkallı, 2014; Taylan, 2015). This works with the consonants /b/, /d/, and /g/ losing their voicing feature and turning into the respective voiceless stops [p], [t], and [k], with an addition of aspiration. The literature for word-final devoicing for other consonants is quite scarce for Turkish. The Turkish rhotic sound was claimed to have word-final devoicing properties by Göksel and Kerslake (2005) and Taylan (2015) while Kornfilt (1997) observed some of the liquid sonorants ([l] and [ʎ]) also show a similar devoicing process. Therefore, this study aims to test the devoicing properties of other voiced consonants in Turkish, namely the [+sonorant] consonants [ɾ], [l], [ɫ], [ʎ], [j]; and [+continuant] voiced fricatives [z], [ʒ], [v], and [β]. We utilized the feature theory in accordance with the works of Chomsky &

* We sincerely thank **Asst. Prof. Stefano Canalis** from Boğaziçi University who has been mentoring this study ever since it was a final paper of a graduate Phonology course. It would not be possible to expand this study to its current state without his support. Authors: Ezgi Sanıyar (ezgi.saniyar@gmail.com) and Ümit Can Tunçer (umitctuncer@gmail.com), Boğaziçi University.

Halle (1997), and Odden (2005) in order to define a feature category for the phonemes and allophones we focus on. This study slowly evolved from smaller sets of data that focused on the aforementioned consonants in separate groups including the nasal stops [n] and [m]. However, these preliminary recordings indicated no devoicing in word-final environment for the nasal stops and therefore were excluded from the list of sonorants. The remaining set of consonants share the feature [+continuant] and create a more uniform group. The results of this study indicate varying amounts of devoicing for these consonants in word-final position. We also suggest that “word-final position” is not an optimal definition for the environment where [+continuant] devoicing occurs. This lends support to Nichols (2016) who draws a line between the stop devoicing and continuant devoicing in Turkish.

2. Methodology. A group of participants was asked to record their voices while reading a dataset that is prepared for this study. 10 participants’ worth of data, 3 male and 7 female, was fully analyzed and used in this study. The participants currently reside in Istanbul with different hometowns across Turkey. All of them have the standard Istanbul accent and they are aged between 20-27. Each participant was asked to read the given words in isolation and in sentences. In total 27 words and 36 sentences with the target continuants in different environments were presented to the participants. They read each item twice. The dataset consists of tokens focusing on the phonemes/allophones [ɾ], [l], [ɫ], [ʎ], [j], [z], [ʒ], [v], and [β] in 7 environments as the following:

a) word-final isolation position
b) sentence-final position
c) pre-voiceless position (__# C[-voi])
d) pre-voiced position (__# C[+voi])
e) pre-vowel position (__# V)
f) voiceless adjacent (__+C[-voi])
g) voiced adjacent position (__+C[+voi])

Any tokens that were unusable due to technical reasons (~5%) were supplemented by the other participants of the study, which were still in the analysis phase. All recordings were converted to WAV formats and made readable for Praat 6.1.52 version. In order to observe the devoicing ratio of sounds, the following parameters and voice reports were extracted from Praat:

i) Total duration of the phone
ii) Duration of the voiced frames
iii) Duration of the voiceless frames
iv) Pulse count
v) Fraction of locally unvoiced frames
vi) Harmonics-to-noise ratio

The total number of tokens that were analyzed was 1260. After the analysis was repeated for each participant, the data was collected in a spreadsheet where the means for each value were calculated. The value of “Fraction of locally unvoiced frames” was taken as the main indicator of devoicing. This value was calculated by Praat which gives slightly different results compared to a manual calculation of voiceless/total duration. While not being an absolute threshold, we considered the tokens with values roughly above 0.15 as partially devoiced. Additionally, the durations of the voiceless and voiced frames were taken as a secondary indicator of devoicing, which constitutes a manual way of calculating devoicing as opposed to the “fraction of locally unvoiced frames" value that is provided by Praat. "Harmonics-to-noise ratio" is a value that correlates with devoicing as well as air friction, which is a helpful side-indicator of devoicing. Lastly, "pulse count" is a value captured by Praat that shows the distribution of regular waveforms, which are produced by the vibration of vocal folds. Its main purpose during the analysis was to determine the borders of voiced and voiceless frames. Although it negatively correlates with the amount of devoicing, interspeaker variation such as the size of the vocal tract (mostly affected by the sex or age of the speaker) makes it difficult to reflect this correlation in a study where the participants are mixed.

3. Results. The results indicate a clear environment of devoicing for the voiced continuant phones of Turkish. In certain environments, [+cont] sounds show a very strong devoicing tendency; this includes word-final position in isolated word and sentence-final position (i.e. utterance final). In the other environments like pre-voiced or voiced adjacent, devoicing ratio is clearly low; for most participants, we don't observe any devoicing tendency for such environments. Below are two screenshots of uttered words in Praat (Figure 1 & 2). The boundaries for the continuant sound /z/ are indicated with the pink color. Figure 1 shows a word where /z/ is in sentence-final position whilst in Figure 2, the /z/ sound is in word-final position preceding a voiced consonant

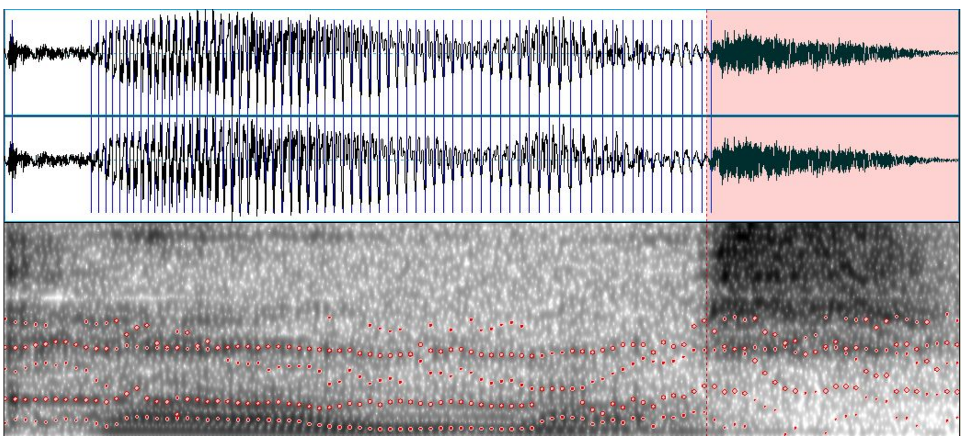

Figure 1. Word /kʰałmaz/ "kalmaz" - *(it) does not stay* in sentence-final position

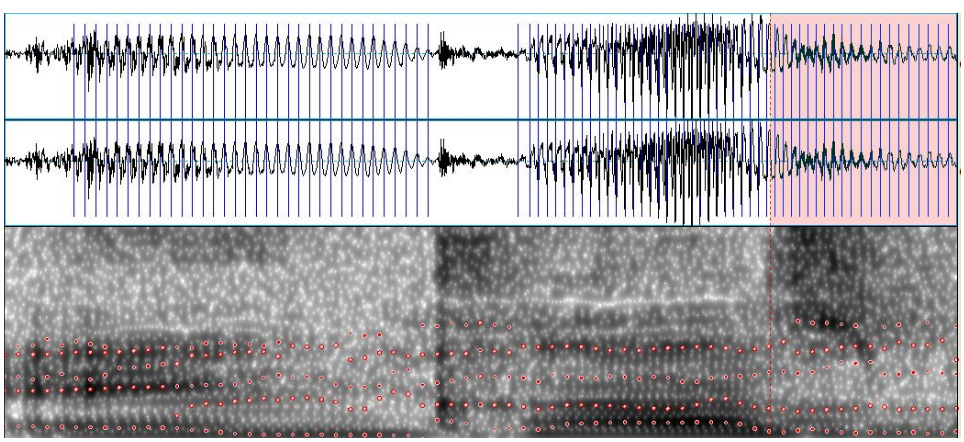

Figure 2. Word /enkʰaz/ "enkaz" - *debris* in __# C[+voi] position

The vertical blue lines, which show periodic pulses, are present during the utterance in Figure 2 while in Figure 1 they are replaced with friction and noise as can be seen from the waveform and the spectrogram, thus no pulse. This shows that in the sentence-final position, the phoneme /z/ lost its voicing quality and was pronounced much like a [s] sound, a voiceless continuant.

The following three tables show the gathered data for three relevant environments to discuss continuant devoicing. Table 1 shows the sentence-final environment which is quite similar to the data in isolation position also analyzed for this study. Together they constitute an "utterance-final position" which we will refer to as a single position from now on. Table 2 shows the data for the word-final position preceding a voiced consonant. This data will be relevant while discussing the general absence of devoicing in word-final position contrary to some findings in the literature. Table 3 shows the data for the syllable-final position preceding a voiceless consonant where another factor, voicing assimilation, could be playing a factor.

Phone	Duration (ms)	Voiced frames (ms)	Voiceless frames (ms)	Pulse count	Fraction of locally unvoiced frames (%)	Harmonics to noise ratio (dB)
[ɾ]	223,55	40,10	183,45	6,55	0,63	4,72
[l]	162,90	70,40	92,50	11,05	0,46	6,25
[ɫ]	160,10	55,20	104,90	8,25	0,50	6,18
[ʎ]	167,95	86,40	81,55	13,10	0,41	7,31
[j]	143,25	62,10	81,15	8,10	0,37	6,56
[z]	241,05	30,15	210,90	4,75	0,81	5,87
[ʒ]	222,65	15,10	207,55	2,25	0,81	3,63
[v]	110,55	51,10	59,45	7,55	0,43	5,76
[β]	156,80	66,70	90,10	10,20	0,40	6,64

Table 1. The phonetic properties of continuant phones in sentence-final position

Phone	Duration (ms)	Voiced frames (ms)	Voiceless frames (ms)	Pulse count	Fraction of locally unvoiced frames (%)	Harmonics to noise ratio (dB)
[ɾ]	83,65	83,65	0,00	17,35	0,01	12,60
[l]	69,30	69,30	0,00	18,25	0,00	17,46
[ɫ]	94,40	94,40	0,00	22,90	0,00	14,40
[ʎ]	86,55	86,55	0,00	23,50	0,00	15,03
[j]	76,50	76,05	0,45	18,10	0,01	12,96
[z]	83,70	79,30	4,40	19,70	0,05	9,80
[ʒ]	105,40	76,90	28,50	18,25	0,18	9,90
[v]	95,20	95,20	0,00	19,90	0,00	11,90
[β]	77,35	77,35	0,00	18,75	0,00	10,86

Table 2. The phonetic properties of continuant phones in word-final position, preceding a voiced consonant (__#C [+voi])

Phone	Duration (ms)	Voiced frames (ms)	Voiceless frames (ms)	Pulse count	Fraction of locally unvoiced frames (%)	Harmonics to noise ratio (dB)
[ɾ]	96,00	62,05	33,95	11,68	0,24	8,76
[l]	99,15	77,15	22,00	14,70	0,18	10,46
[ɫ]	112,15	94,90	17,25	18,65	0,12	14,14
[ʎ]	70,25	69,25	1,00	13,30	0,02	13,11
[j]	97,05	85,05	12,00	16,80	0,08	10,47
[z]	143,05	29,75	113,30	5,75	0,76	6,36
[ʒ]	106,70	38,60	68,10	7,95	0,56	8,68
[v]	71,40	64,20	7,20	11,90	0,08	10,73
[β]	86,95	78,95	8,00	15,35	0,07	13,00

Table 3. The phonetic properties of continuant phones in syllable-final position, preceding a voiceless consonant (__+C [-voi])

The tables above further show that devoicing ratios of the continuants differ substantially and while some of them show a clear devoicing tendency in utterance-final position (Table 1), some of them have lesser or optional devoicing. Among them, especially fricative sounds like [z] and [ʒ] and rhotic [ɾ] have the strongest devoicing amounts and tendency, while the sonorants [ʎ], [ɫ], [l] and the fricatives [v] and [β] have increased participant-based differences and optionality in utterance-final devoicing. Mean values of "fractions of locally unvoiced frames" (shown in green) in sentence-final and isolation environments draw the following order (Example 1) of devoicing intensity among the phones:

(1) [z] > [ʒ] > [ɾ]> [l]> [ɫ]> [β] > [v] > [j] > [ʎ]

Comparative values of pulse counts and harmonics-to-noise ratios also indicate a devoicing process, favoring the utterance-final environments. In Table 2, the values of fraction of voiceless frames drop significantly for [ɾ], [j], [z], [ʒ] and the other sounds exhibit no devoicing at all even though it is technically a word-final position. This is a further evidence for that the word-finalness is not an optimal environment to attribute the devoicing phenomenon for these sounds. Table 3 marks an interesting phenomenon where fricative sounds have significantly greater amounts of devoicing in syllable-final position preceding voiceless consonants than in word-final position preceding voiced consonants. This might indicate that voicing assimilation can also be playing a role in continuant devoicing. The continuant sounds showed similar amounts of devoicing in word-final position preceding a voiceless consonant (Table 4) which strengthens the argument for the existence of voicing assimilation.

Phone	Duration (ms)	Voiced frames (ms)	Voiceless frames (ms)	Pulse count	Fraction of locally unvoiced frames (%)	Harmonics to noise ratio (dB)
[ɾ]	84,95	67,75	17,20	17,55	0,13	8,27
[l]	87,90	73,52	14,38	20,24	0,09	11,47
[ɫ]	83,05	78,35	4,70	20,40	0,06	12,61
[ʎ]	101,30	83,20	18,10	22,55	0,13	10,79
[j]	81,60	78,75	2,85	19,90	0,02	10,57
[z]	99,25	35,35	63,90	8,45	0,56	5,96
[ʒ]	107,50	23,30	84,20	5,80	0,71	5,25
[v]	74,40	59,95	14,45	15,40	0,14	8,52
[β]	72,85	65,25	7,60	15,55	0,10	10,67

Table 4. The phonetic properties of continuant phones in word-final position, preceding a voiceless consonant (__#C [-voi])

Though rare, some participants exhibited different voicing qualities for the same phone in the same environment. We suspect that the phonological gaps between two tokens and reading pace have also some effect on [+cont] devoicing. Our experiment did not strictly control the reading pace enough to produce reliable data we can argue for this variable. However, we can speculate that a greater phonological gap is likely to be the main factor of devoicing in utterance-final positions.

Lastly, Table 5 shows the standard deviation values for the “Fraction of locally unvoiced frames” values across the participants. Compared with the average values of their respective environments in the previous tables, these values show the variation in optionality and devoicing amounts. While in sentence-final position the fricatives [z] and [ʒ] show lesser deviation, favoring lesser optionality for voicing and longer devoiced fractions, the values increase in non-utterance-final environments. We can also see that the rhotic, while having larger fractions among other sonorants on average, shows a similar standard deviation to the other sonorants. This indicates that, among the sonorants, the rhotic is more likely to be devoiced in addition to its highest devoiced fractions. Pre-voiceless (__# C[-voi]) and voiced adjacent (__+C[+voi]) positions, where we previously mentioned a possible role of voicing assimilation, shows the deviation coming close to and often surpassing the mean value of voiceless fractions, suggesting that these environments have the highest inter-speaker variance in terms of devoicing optionality.

Phone	Sentence-Final	__#C [+voi]	__+C [-voi]	__#C [-voi]
[r]	0,20	0,02	0,24	0,14
[l]	0,23	0,00	0,17	0,11
[ɫ]	0,22	0,00	0,18	0,07
[ʎ]	0,19	0,00	0,04	0,16
[j]	0,20	0,04	0,14	0,04
[z]	0,08	0,14	0,19	0,25
[ʒ]	0,13	0,31	0,25	0,18
[v]	0,23	0,00	0,12	0,14
[β]	0,25	0,00	0,13	0,13

Table 5. Standard deviation values of "Fraction of unvoiced frames" across participants

4. Conclusion. The data collected for this research provides a piece of strong evidence for the devoicing properties of [+cont] consonants in Turkish. While different positions have an effect on devoicing ratio in continuants, utterance-final position is clearly the most prominent environment for devoicing which separates this from the stop devoicing process in Turkish, which is observed in every word-final position. That being said, we also suggest that the previous descriptions of devoicing in the literature for the rhotic (Göksel & Kerslake, 2005) and liquids (Kornfilt, 1997) become more consistent if the descriptive environment is shifted from word-final to utterance-final, as Nichols (2016) also suggests for the rhotic. The optionality and intra-speaker variation also indicate that this is a different kind of devoicing process. We do not observe stable and equally detectable devoicing in [+cont] as we do in the stop devoicing process. For instance, depending on the participant's pronunciation, a continuant may or may not show devoicing properties in pre-voiceless or voiceless adjacent positions. Although this optionality is quite limited and the findings are mostly consistent, we may observe the same differences in final positions, as well.

As the data shows, fricatives are the most sensitive ones in this process, some sonorants like the rhotic consonant has also a high ratio of devoicing in the final position, as previously observed by Göksel & Kerslake (2005). With the data collected and all the findings, we could claim that devoicing is not only limited to [-cont, -son] sounds in Turkish and this process is clearly observable in other [+cont] consonants, though they are not strictly and equally applicable for all and some may have optionality to a certain extent.

5. Implications for further research. For future work, a study can be conducted to investigate the effects of the adjacent or pre [+/-cont and +/-son] sounds on the ratio change, different syntactic positions, and related varying phonological gaps. Because some of the devoicing was observed preceding voiceless consonants, the whole group of voiced continuants also poses a potential research area for voicing assimilation processes in Turkish. One environment that was left out from this study is word-final CC clusters as it is quite limited for some of these continuants (e.g. */zt/, */zd/, */ʒt/ etc.). However, the ones that are available for a CC cluster (e.g. /rt/, /vt/, /lt/ etc.) can be studied as an additional environment for voicing assimilation.

Additionally, we believe that different dialects, accents, sociological status, gender/age differences, and user perception are important factors that may have an effect on participants' devoicing tendencies. Recordings of a spontaneous speech rather than the speech in a controlled experimental environment can also give different results and discussion areas. As the dataset is very much tailored and the utterances are collected by a reading process, observing how continuant devoicing in casual speech work could be an interesting follow-up to this study. In a further study, these can be taken into consideration and comparable results can be discussed.

References

Chomsky, N., & Halle, M. (1997). The sound pattern of English. The MIT press

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

Kopkallı, H. (2014). A phonetic and phonological analysis of final devoicing in Turkish. Anadolu University (Turkey).

Kornfilt, Jaklin. 1997. Turkish grammar. London, Routledge.

Nichols, Stephen. 2016. An acoustic study of the Turkish rhotic. Poster presented at the 5th International Workshop on Phonetic, Phonological, Acquisitional, Sociolinguistic and Dialect Geographic Aspects of Rhotics, Leeuwarden 18-20 May 2016.

Odden. (2005). Introducing Phonology (Cambridge introductions to language and linguistics). Cambridge University Press.

Van der Hulst, H., & Van De Weijer, J. (1991). Topics in Turkish phonology. Turkish linguistics today, 11-59.