PROSODIC AND SEGMENTAL INFLUENCES ON HIGH VOWEL DEVOICING IN TURKISH

Stefanie Jannedy

Department of Linguistics
Ohio State University

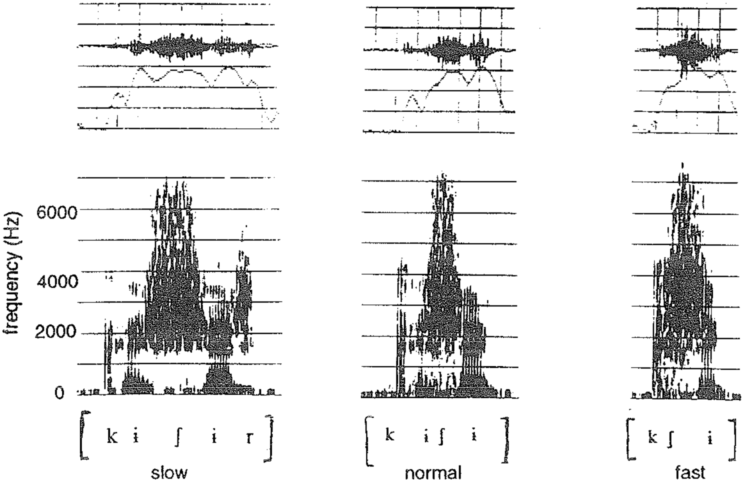
ABSTRACT - Prosodic and segmental factors such as *rate, stress, preceding environment, following environment, vowel-* and *syllable type* which influence the process of vowel devoicing in Turkish are described and evaluated. Results are contrasted with findings for Japanese and Korean. Browman & Goldstein's Gestural Score Model (1990) is employed to explain the data.

INTRODUCTION

A previously undocumented process of vowel devoicing affects the four high vowels /i y ɨ ɯ/ of Turkish. Previous studies of Japanese and Korean have shown that the change in specification from [+voice] to [-voice] is not categorical but rather gradual, there are intermediate stages found on a continuum from a fully voiced to a completely deleted vowel, where a partially devoiced vowel has only a few very weak glottal pulses. Jun & Beckman (1993, 1994) thus propose to analyze vowel devoicing in terms of gradual gestural overlap (Browman and Goldstein, 1990) where the glottal gestures for preceding and following voiceless consonants overlap to a greater or lesser extent with the [+voice] glottal gesture for the high vowels. The devoicing of the four short high Turkish vowels /i y ɨ ɯ/ can be explained by the same model that predicts that vowels are more likely to undergo devoicing if they are short and the adjacent voiceless consonants have large glottal opening gestures.

METHODS

Nine native speakers of Turkish read 171 words positioned utterance initially in carrier-phrases at three rates (slow, normal, fast). Each of the 4617 tokens was rated as containing either a voiced (clear voice bar with several glottal pulses), partially devoiced (one or two faint glottal pulses) or fully devoiced (no glottal pulses visible at all) vowel. The criteria for this categorization on a voicing continuum are similar to the ones used by Jun and Beckman (1994) for Korean.



slow normal fast

Figure 1. Spectrograms, waveforms and rms amplitude traces of voiced, partially devoiced and completely devoiced vowel tokens in the word [kɨʃɨr] 'crust, bark'

Varbrul (Variable Rule) analysis (Sankoff, 1988; Rand & Sankoff, 1990) was used to evaluate the relative importance to the distribution of devoicing of the different predictors such as *rate, stress, preceding-* and *following environment, vowel-* and *syllable type*. This analysis method uses step-wise multiple regressions on a logistic transform of the proportions of tokens which undergo a "rule" (in this case, vowel devoicing) for each combination of factors, with a maximum likelihood estimation criterion to accommodate imbalances of number of tokens within the various cells.

RESULTS

All prosodic and segmental factors were significant, based on an adjusted alpha of $p < .002$. High vowels where more frequently devoiced at faster speaking rates and in unstressed syllables. Stops and fricatives (both 16% completely devoiced tokens) and affricates (26%) in the preceding environment were more closely associated with the process than no preceding environment (9%). Ranked by consonant type, the relative strength of influence was [tʃ] (26%) > [p] (19%) > [k] (16%) > [f] (15%) > [ʃ] = t (13%) > [s] (12 %) > zero-context (9%). As for the influence of the following environment, overall devoicing was more likely to occur with following stops (17%) than fricatives (12%). The individual ranking is as follows: [t] = [k] (19%) > [tʃ] (15%) > [s] (14%) > [p] = [ʃ] (12%) > [f] (8%). Vowel quality was a significant factor in the Variable Rule analysis of vowel devoicing whereby [i ɨ y] were devoiced more often than [u]. Considerably less devoicing occurred in closed syllables (10%) compared to open syllables (21%).

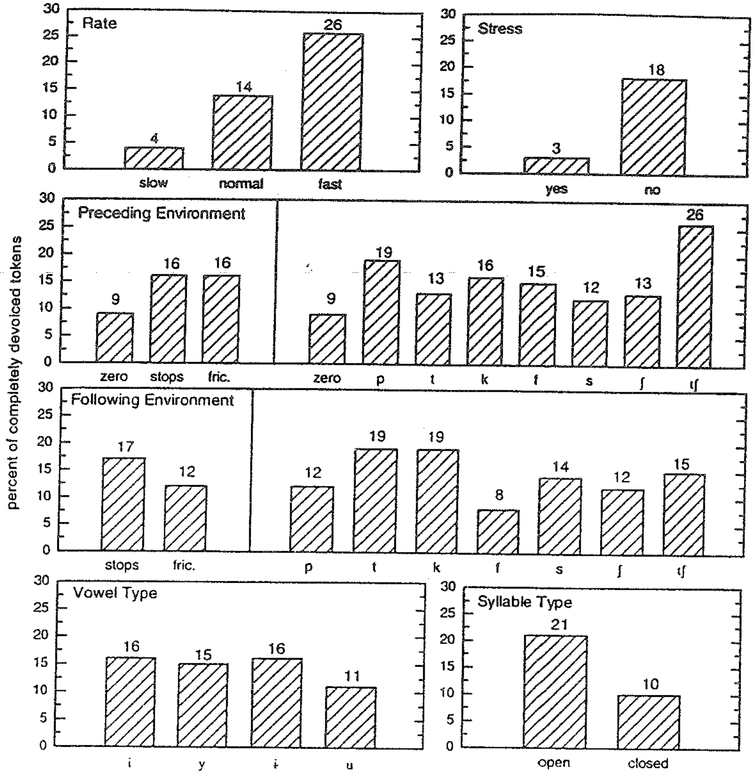


Figure 2. Influence of rate, stress, segmental environment, vowel- and syllable type

To evaluate the influence of the preceding environment and of syllable type, further analyses were conducted to establish the mean voice onset time (VOT) for Turkish voiceless stops before non high vowels and the duration of non-high vowels in open and closed syllables. Five speakers read 28 pairs of words in two conditions: three times in different randomized order in isolation and utterance initially positioned in a carrier-phrase in three different speech rates (slow, normal, fast). VOT of word initial stops was measured from the release of the stop burst to the onset of voicing, visible as a voice bar on the spectrogram on a KAY DSP-5500. Measurements were also taken for the duration of non-high vowels from the onset to the end of the vowel's formant structure in the initial unstressed syllable.

Mean VOT-values for Turkish unaspirated voiceless stops before non-high vowels in a normal speech rate in citation were 48 ms in open and 49 ms in closed syllables and not significantly different. These durations are slightly longer than mean values cited in Lisker & Abramson (1964) for other languages with unaspirated, but shorter than for aspirated voiceless stops. Also, according to measurements of the duration of non-high vowels in open and closed syllables of two sets of words, vowels are, contrary to findings for many other languages (Maddieson, 1985) significantly longer in closed syllables than in open syllables. This result was confirmed in a replication of an experiment (15[1] [minimal]-pairs of words with single consonants and geminate consonants closing the syllable) done by Lahiri & Hankamer (1988) who found non-significantly longer vowels in closed syllables, and with 28 other open/closed-syllable minimal pairs. All words were elicited in isolation (three repetitions) as well as in initial position in a carrier phrase (in three different rates). Measurements were made according the criteria described above. Also, since Lahiri and Hankamer originally measured from waveforms and not spectrograms, measurements for five pairs of their words were repeated from waveforms and correlated with the measurements for the identical token obtained from spectrograms. The correlation was $r = .82$ for open and $r = .85$ for closed syllables. A paired t-test showed that even with a relatively small n of 75, mean differences in vowel duration in open and closed syllables were highly significant regardless of measurement tool (spectrogram: $t = -8.84$, $p < .001$; waveform: $t = -6.86$; $p < .001$). The graph shows data from pairs of words repeated three times in citation form ($n = 419$, and 222; in total, four token had to be discarded). Paired and independent t-tests showed significant differences on $p < .001$. Mean duration of vowels as measured from spectrograms of 28 pairs of words used in this ('94) study were: closed syllable: 77 ms; open syllable: 66 ms; 15 pairs from Lahiri & Hankamer's 1988 study: closed syllable: 87 ms, open syllable: 78 ms.

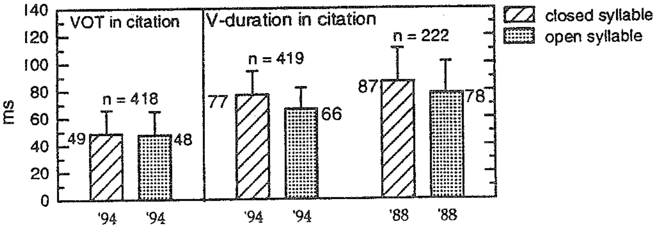


Figure 3. Mean duration of VOT in initial syllables (left) mean duration of non-high vowels in open and closed syllables (right).

DISCUSSION

Rate

Munhall and Loefqvist (1992) elicited multiple renditions of the phrase *Kiss Ted* in different speech rates, ranging from slow to fast. In slower renditions they found two distinct glottal opening and closing movements, one for the [s] in *Kiss* and another for the [t] in *Ted*. With an increasing rate of speech, these two glottal gestures began to overlap and in the fastest rates they were completely blended into one glottal opening and closing gesture. Based on these results, faster tempi should trigger more devoicing of high vowels than slower ones since the preceding and the following voiceless consonants' glottal gesture are assumed to overlap with the vowel's glottal gesture. The gestural score model (Browman & Goldstein, 1990) predicts that the vocalic and consonantal gestures have greater stiffness the faster the rate of speech because the articulatory movements are executed faster. The finding that an increased rate had the most impact on the process of gradual vowel devoicing can be explained similarly, that is, glottal gestures of adjacent consonants overlap the vowel's voicing specification. Predictions of the model appear to be borne out well, devoicing of vowels occurred more often at faster rates because the glottal gesture for the vowel has less time allotted for reaching its target (regular vocal fold vibration), and thus phonetic undershoot (Lindblom, 1969) can be caused by faster rates of speech.

Stress

In languages like English, unstressed vowels are more likely to undergo devoicing than stressed vowels because unstressed vowels are shorter and more centralized and thus more prone to reduction. The first syllables in *potato* and *Chicago* are such examples. If stress in Turkish is comparable to stress in English, we can predict that unstressed vowels are more likely to undergo vowel devoicing because their voicing gesture is shorter and thus more easily overlapped by adjacent consonantal laryngeal gestures. As the graph in Figure 2 shows, these predictions hold true for Turkish, more devoicing is found for unstressed syllables, where vowels are shorter.

Segmental Environment

In Turkish devoicing also occurred in utterance-initial position, thus creating syllable-initial consonant clusters like *sp-* or *st-* which according to Kornfilt (1987) are not allowed in Turkish. Examples are /ispirto/ -> [spirto] 'alcohol' or /istampa/ -> [stampa] 'inkpad'. Van der Hulst & van de Weijer (1991) treat initial high unrounded vowels in Turkish as inserted by prothesis for the purpose of breaking up and resyllabifying the consonant cluster. Thus, where the initial high vowels are devoiced, the phonological process is violated by the phonetics. Utterance initial vowel devoicing is less easy to interpret in terms of the gestural score model, because only the following laryngeal gesture can overlap with the vocalic gesture and thus, to achieve complete devoicing, must be fairly strong.

According to the gestural overlap model, most devoicing should be found after consonants with large glottal opening gestures. For Korean, Jun & Beckman (1994) find slightly more devoicing after plain fricatives than after aspirated stops. This finding accords well with Kayaga's findings (1974) of generally longer sustained peak glottal opening for fricatives compared to stops. Cedergren and Simoneau (1985) find generally more devoicing after voiceless fricatives than after voiceless unaspirated stops in Montreal French. For Japanese, Nagano-Madsen (1994) reports on research conducted by other scientists that also shows more vowel devoicing after fricatives although Japanese data provided by Sawashima and Hirose (1983:17) shows no difference in VOT (onset of voicing) after a voiceless fricative or unaspirated stop.

In Turkish, most devoicing is found after affricates which, preceding a vowel should behave like fricatives since the right edge of an affricate is a fricative. However, affricates appear to act more like stops when preceding vowels since generally more devoicing is found after stops than fricatives. The difference in behavior between fricatives and stops is obscured by the mean devoicing rates since the high rate for the affricates (grouped with the fricatives) averages out the actual difference between fricatives and stops. Thus, if affricates are grouped with stops, then devoicing actually occurs more frequently following stops. Since these results do not confirm findings made for Korean, Japanese or French where more devoicing is found after fricatives which (in Korean and Japanese) have a longer peak glottal opening, duration of peak glottal opening during consonant articulation or possibly the size of the glottal opening might be language specific. One could speculate that peak glottal opening during fricative articulation occurs so early that it does not extend into the oral gesture for the vowel. Also, one might assume that the size and not phasing or duration of the glottal opening is relevant since affricates are most influential. Air pressure initially builds up before the stop part of the affricate. As the air is released into a fricative, the vocal folds get pushed open widest so that the air can escape to release the stop. Even though airflow must be sustained to cause frication at some obstruction in the oral cavity, the initial release size and force of the stop might be an explanation for the patterning of the affricates with the unaspirated stops.

As for the impact of the following environment on vowel devoicing, results for Turkish are consistent with findings of Jun & Beckman (1994) for Korean where more devoicing was found before stops than fricatives. As data by Sawashima and Hirose (1983) show for Japanese, vocal fold vibration ceases sooner in vowel-stop sequences than before fricatives since airflow through the glottis ends abruptly. However, Cedergren and Simoneau (1985) report generally more devoicing before voiceless fricatives than before affricates and stops in Montreal French.

Vowel Type

All four Turkish high vowels can undergo the process of gradual vowel devoicing, however [u] is slightly more resistant to devoicing than [i ɨ y]. This resistance of [u] is difficult to explain, although a study of vowel duration in Turkish might show that [u] is longer than the other high vowels in Turkish.

Syllable Type

Even though the factor syllable type has least impact on the process of vowel devoicing, it significantly contributed to increase the likelihood of vowel devoicing. Contrary to the expectations (based on data for other languages), vowel duration is significantly shorter in open syllables ($p < .001$) than in closed ones. Thus, since more devoicing is expected for shorter vowels, the outcome is just as we might expect. The durational differences of vowels in open and closed syllables is a fairly unusual observation and difficult to explain. Whether the statistically significant difference in duration of vowels in production is perceptually important and contrasting, is still yet to be shown. However, the very robust effect of finding longer vowels in closed syllables through different experimental conditions (read in different speech rates, within carrier phrases, in isolation, measurements from spectrogram and waveform) is interesting in itself. For once, the data suggests that vowel duration in open and closed syllables is not an articulatory artifact of human speech production or some kind of rhythmic constraint, but that vowel duration is language specific and independent of the syllable structure. Han (1994) mentions the possibility of pre-geminate lengthening to explain the 11% longer segment duration before geminates in Japanese. However, in Turkish longer vowels were found in closed syllables before geminates and also before various other consonant combinations, C_1 being the coda of the first syllable and C_2 being the onset of the second one. For Turkish this could mean that vocalic gestures in closed syllables have a later onset phase target with regard to the preceding consonant than within open syllables. Phonotactic constraints on the possible shape of a syllable are potentially important as well: the complete devoicing of a syllable initial vowel in words like *istanbul* 'istanbul (city)' or *ispanak* 'spinach' generates undesired and less frequent syllable onset clusters like *st-* or *sp-* (attested in some loan words), and thus complete devoicing or deletion of the vowel in this position is possible but relatively rare. Vowels might be longer in closed syllables with a $C_1VC_2.C_3$ structure so that syllable structures like $C_1C_2.C_3$ or $C_1.C_2C_3$ are prevented after devoicing or deletion and resyllabification.

CONCLUSIONS

The presented data showed that various prosodic (rate, syllable type, stress) and segmental (null-context, preceding environment, following environment, vowel type) factors influence the process of high vowel devoicing in Turkish. Generally, intrinsically shorter high vowels devoiced more easily when they were weakened further by prosodic factors (fast rate, lack of stress) or shorter due to the environment. To a large extent, the findings could be explained by the gestural overlap model which accounts for the influence of the individual factors in terms of overlap and blending of glottal gestures of adjacent voiceless consonants with the vocalic glottal gesture. The data also possibly suggests language specific intersegmental timing relations between consonantal and vocalic oral-gestures.

REFERENCES

Browman, C. & Goldstein, L. (1990) *Tiers in articulatory phonology, with some implications for casual speech.* Kingston, J. & Beckman, M. E. (eds.) Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech, 341-376. Cambridge: Cambridge University Press.

Cedergren, H. J. & Simeneau, L. (1985) *La chute des voyelles hautes en français de Montreal <<As-tu entendu la belle syncope?>>.* Lemieux, M. & Cedergren, H. J. (eds.) Les tendences dynamiques du francais parle a Montreal, Tome 1, 57-144.

Han, M. S. (1994) *Acoustic Manifestations of mora timing in Japanese.* Journal of the Acoustical Society of America, 96, No. 1, 73- 82.

Jun, S.-A. & Beckman, M. E. (1993) *A gestural-overlap analysis of vowel devoicing in Japanese and Korean.* Paper presented at the 1993 Annual Meeting of the Linguistic Society of America, 7-10 January 1993, Los Angeles, CA, USA.

Jun, S.-A. & Beckman, M. E. (1994) *Distribution of Devoiced High Vowels in Korean.* Proceedings of the International Conference on Spoken Language Processing, Yokohama, Japan, Vol. 2, 479-482.

Kayaga, R. (1974) *A Fiberscopic and Acoustic Study of the Korean Stops, Affricates and Fricatives.* Journal of Phonetics, Vol.2, 161-180.

Kornfilt, J. (1987) *Turkish and the Turkic Languages.* Bernhard Comrie (ed.) The World's Major Languages, 619-644. Croom Helm Ltd.: London, Sydney.

Lahiri, A. & Hankamer, J. (1988) *The Timing of Geminate Consonants.* Journal of Phonetics,16, 327-338.

Maddieson, I. (1985) *Phonetic Cues to Syllabification.* Fromkin, V. A. (ed.) Phonetic Linguistics. Essays in Honor of Peter Ladefoged, ch. 14, 203-221.

Munhall, K. & Loefqvist, A. (1992) *Gestural Aggregation in Speech: laryngeal gestures.* Journal of Phonetics, 20, 111-126.

Nagano-Madsen, Y. (1994) *Vowel Devoicing Rates in Japanese from a Sentence Corpus.* Department of Linguistics and Phonetics, Lund, Sweden, Working Papers 42, 117-127.

Rand, D. & Sankoff, D. (1990) *Gold Varb, Version 2, A Variable Rule Application for the Macintosh.* Centre de recherches mathematiques Universite de Montreal.

Sankoff, D. (1988) *Variable Rules.* Ammon, U., Dittmar, N., Mattheier K. J. (eds.) Sociolinguistics: An International Handbook of Science and Language, Vol. 2, ch. 112, 984-997.

Sawashima, M. & Hirose, H. (1983) *Laryngeal Gestures in Speech Production.* MacNeilage, P. F. (ed.) The Production of Speech, ch. 2, 11-38. Springer Verlag, New York, Heidelberg, Berlin.

Van der Hulst, H.; Van de Weijer, J. (1991) *Topics in Turkish Phonology.* Boeschoten, H. & Verhoeven, L. (eds.) Turkish Linguistics Today, 11-59. E. J. Brill, Leiden, The Netherlands.

ACKNOWLEDGMENTS

This work has benefited much from comments and discussions with Mary Beckman, Marios Fourakis, Keith Johnson, Sun-Ah Jun and Jennifer Venditti. The Department of Linguistics at the Ohio State University is thanked for their generous travel- and consultant support.

---

[1] Lahiri and Hankamer's 1988 experiment tested 18 pairs of words: Two pairs were discarded in this study because of difficulties finding consisten measurment landmarks. The third pair (saate - saattee 'clock' DAT and LOC) was analyzed as a long vowel but all five informants for this study analyzed the vowel sequence as having a syllable break in the middle.