Tongue root position in VC sequences with regard to the phonetic realisation of obstruent voicing: A preliminary study on Hungarian

Gráczi, Tekla Etelka[1,5] – Csapó, Tamás Gábor[2,5] – Deme, Andrea[3,5] – Juhász, Kornélia[4,5] – Markó, Alexandra[3,5]

[1]Hungarian Research Institute for Linguistics, [2]Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics, [3]Eötvös Loránd University, Department of Applied Linguistics and Phonetics, [4] Eötvös Loránd University, Doctoral School of Linguistics, , [5]MTA-ELTE "Lendület" Lingual Articulation Research Group

Corresponding author: graczi.tekla.etelka@nytud.hu

Introduction. The simultaneous articulation of the turbulent noise of fricatives and vocal fold vibration has conflicting pressure requirements. Among various differences between voicing counterparts, the tongue root was found to be advanced in the voiced segments in various languages compared to the voiceless counterparts (e.g., [1-4]). This difference appeared already in the preceding vowel before stops. The maintenance of vocal fold vibration in voiced fricatives were found to vary not only across speakers [5] but also across languages [6]. The articulatory manoeuvres also showed differences between languages and speakers: tongue root advancement was found in Italian stops at closure onset and maximum displacement but not in Polish [4]. The tongue-palate contact was strongly related to the amount of voicing present in the fricatives in devoicing speakers [5].

As the phonetic realisation of voicing varies across speakers (for Hungarian see e.g. [7]), our question is, what patterns tongue root advancement shows in the obstruents and in its preceding vowel with regard the obstruents' voicing patterns in Hungarian. The present experiment introduces a pilot study comparing /iz/ and /is/ realisations. Our hypothesis was that speakers whose /z/ realisations tend to maintain voicing throughout the (most of) its duration have larger differences in the fricative and the preceding vowel between /z/ and /s/ than devoicing speakers.

Methods. 12 female, monolingual native speakers of Hungarian without any speech or hearing disorder read aloud 5 sentences started with /izi nɛ/ and 5 with /isi nɛ/. The stimuli of the present pilot study were embedded into another study material [8]. The target sounds of this experiment were the /z/ and /s/ and the preceding /i/. These /i/ realisations bore both sentence accent and word-stress.

The tongue contours (TC) were recorded in midsagittal orientation using the "Micro" ultrasound (US) system (AAA, Articulate Instruments Ltd). The vocal fold activity was captured by a D200 EGG (Laryngograph LtD.). The speech signal was recorded with a head-mounted microphone (time aligned to the ultrasound recording automatically by the recording software) and with a clipped microphone (at fix distance from the mouth, recorded by the EGG). The EGG- and US-recordings were time aligned through the two speech signals. The segmentation was carried out by forced alignment [9] and manually corrected in Praat [13].

The TCs of the target sounds were traced manually in AAA at each frame along the time course of the V. The ratio of the entire duration of vocal fold vibration in the C to the C's entire duration (voiced part ratio: VPR) was automatically labelled and manually corrected in the EGG-signal in Praat [10]. Partial devoicing was defined as at least 25% devoicing in the C.

Generalised additive models (GAM) were used in R [11] with the means of the packages mgcv [12], itsadug [13] and rticulate [14]. The TCs were analysed separately for each speaker [see 14]. Two models were built: one for the vowel TC in the V and one for the TC in the VC sequence, where the V was set to the 0-0.5 time interval, the C to the 0.5-1 interval.

A basic model without C contrast and a model with C contrast (with contrast treatment) were compared with χ^2 test. The results showed that the model with C contrast described the data significantly better in all polar GAMs. Therefore, autocorrelation correction (first order autoregressive model) and knot number setting were carried out on these models. The models included the position of the TC points as dependent variable, the smooth of the point of the TC and the normalized time course in the V or VC sequence, and their interaction (tensor). All the smooths and tensors were included with and without nesting for the difference smooths of the C.

Results. Three speaker groups were defined based on VPR: 'voicing' group (0-1 realisations were partially devoiced in 6 speakers' speech: sp01, sp02, sp03, sp04, sp11, sp12), 'partially devoicing' group (2-3 realisations were partially devoiced in 3 speakers' speech: sp05, sp07, sp08), and 'devoicing' group (4-5 realisations were devoiced in 3 speakers' speech: sp06, sp09, sp10).

The models for TC position in V with C contrast explained at least the 90.9% of the deviance of the data in each speaker. There was no general difference in the position of the TC: the results of the parametric terms are significant only in 4 out of the 12 speakers. However, the TCs' change along the time course of the V was significantly different between the /i/ realisations in all speakers.

The models for TC position in VC sequences with C contrast explained at least the 90.4% of the deviance of the data in each speaker. General difference was found in 7 speakers in the TC between the two VC types. The tensor expression gave significant results in 9 speakers, the TC smooth in 9 speakers, the normalised time course smooth in 3 speakers.

In the ‘voicing’ group, sp04 did not show any difference in the tongue root (TR) between the two sequences, while sp01 did but it appeared only in the C. The TR advancement appeared earlier, during the V in the further 4 speakers’ speech (sp02, sp03, sp11, sp12). In the ‘partially devoicing’ group, the TR was advanced in /iz/ realisations in sp05 and sp07, and it appeared already during the V. Sp08 did not show advanced TR in /z/s. In the ‘devoicing group’, the TR was advanced in all three speakers in /iz/ realisations at least from roughly the mid of the V.

Discussion and conclusions. The tongue contour differences between /iz/ and /is/ sequences showed interspeaker differences, however, the presence or absence and the start of the TR advancement did not follow the voicing tendencies of the speakers. Two speakers did not show advanced TR in /z/s at all, and it appeared only later, during the fricative in one further speaker’s speech. As TR advancement is not the only articulatory manoeuvre to maintain voicing, further analyses will be carried out with regard the position of the anterior parts of the TC and the glottal periods.

References

[1] Narayanan, S. S., Alwan, A. A. & Haker, K. (1995). An articulatory study of fricative consonants using magnetic resonance imaging. JASA, 98, 1325–1347.

[2] Westbury, J. R. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing. JASA, 73(4), 1322–1336.

[3] Ahn, S. (2018). The role of tongue position in laryngeal contrasts: An ultrasound study of English and Brazilian Portuguese. J Phonetics, 71, 451–467. doi:10.1016/j.602wocn.2018.10.003.

[4] Coretta, S. (in press). Longer vowel duration correlates with greater tongue root advancement at vowel offset: Acoustic and articulatory data from Italian and Polish. JASA. Preprint downloaded from https://osf.io/zrqyx.

[5] Fuchs, S. Brunner, J. & Busler, A. (2007). Temporal and spatial aspects concerning the realizations of the voicing contrast in German alveolar and postalveolar fricatives. Advances in Speech–Language Pathology, 9(1), 90–100.

[6] Shih, C., Möbius, B. & Narasimhan, B. (1999). Contextual effects on consonant voicing profiles: A cross-linguistic study. Proc. 14th ICPhS. 989–992.

[7] Gráczi, T. E. (2012). Zörejhangok akusztikai fonetikai vizsgálata a zöngésségi oppozíció függvényében. [Acoustic characteristics of obstruents with regard to voicing opposition]. PhD-thesis. Budapest: Eötvös Loránd University. http://doktori.btk.elte.hu/lingv/graczitekla/diss.pdf

[8] Markó, A., Bartók, M., Csapó, T. G., Deme, A. & Gráczi, T. E. (2019). The effect of focal accent on vowels in Hungarian: articulatory and acoustic data. In: Calhoun, S., Escudero, P., Proc. 19th ICPhS. (pp. 2715–2719). http://intro2psycholing.net/ICPhS/papers/ICPhS_2764.pdf

[9] Mihajlik, P., Tüske, Z., Tarján, B., Németh, B. & Fegyó, T. (2010). Improved recognition of spontaneous Hungarian speech: Morphological and acoustic modeling techniques for a less resourced task. IEEE Transactions on Audio, Speech and Language Processing, 18(6), 1588–1600.

[10] Boersma, P. & Weenink, D. (2019). Praat: doing phonetics by computer. http://www.fon.hum.uva.nl/praat/download_win.html. (Downloaded: 2019. october 1.)

[11] R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[12] Wood, S. N. (2017). Generalized Additive Models: An Introduction with R (2nd edition). Chapman and Hall/CRC.

[13] van Rij, J., Wieling, M., Baaye, R. & van Rijn, H. (2017). itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs. R package version 2.3.

[14] Coretta, S. (2019). rticulate: Ultrasound Tongue Imaging in R. R package version 1.5.0.9000. https://github.com/stefanocoretta/rticulate

Acknowledgement

The authors are thankful to Valéria Krepsz for her help with the annotation. The study was supported by the Bolyai János Research Scholarship of the Hungarian Academy of Sciences, the ÚNKP-19-4 New National Excellence Program of the Ministry for Innovation and Technology, and Thematic Excellence Program of the Ministry for Innovation and Technology, Hungary.