Eötvös Loránd University
Faculty of Arts

PhD thesis

Tekla Etelka Gráczi

Zörejhangok akusztikai fonetikai vizsgálata a zöngésségi oppozíció függvényében
Acoustic phonetic analysis of obstruents with regard to voicing contrast
Abstract booklet

Doctoral School of Linguistics
Prof. Dr. Vilmos Bárdosi, head of doctoral school
Program of Hungarian Linguistics
Prof. Dr. Jenő Kiss, head of program

PhD-comittee:
Chairman: Prof. Dr. István Nyomárkay, professor emeritus
Reviwers: Prof. Dr. Klára Vicsi DSc.
Dr. Koloman Brenner PhD.
Members: Dr. Ildikó Bodnár CSc.
Dr. Árpád Zimányi CSc.
Additional members: Dr. György Szaszák PhD.
Dr. Krisztina Menyhárt PhD.

Supervisor: Prof. Dr. Mária Gósy DSc.

Budapest, 2012

Introduction

The obstruents in Hungarian can be considered symmetrical with respect to voicing contrast, as it can be found at all places of articulation of stops (both plosives and affricates) and fricatives. [1]

However the articulatory, aerodynamic characteristics of obstruents result in difficulties in maintaining phonation, therefore vocal cord vibration may cease during their pronunciation (Stevens 1998). The reason is that there is no airflow (closure) or only a small amount of airflow (stricture) from the mouth during the constricted phase that causes intraoral, thus supraglottal, pressure build-up, and as a consequence the necessary condition of phonations, i.e. the transglottal pressure drop, decreases and may be lost. The resonance of the vocal cords is and can be helped by passive and active articulatory gestures (e.g. Stevens 1998; Maddieson 1997; Ohala 1997, Shadle 1997). The passive ones are the consequences of some of the aerodynamic or articulatory features (like supraglottal cavity enlargement due to the same pressure build-up, or lowering of the larynx for phonation). Active compensatory gestures are not the result of other features, but are optional (e.g. lowering of the radix, leakage during closure, nasal cavity opening). Passive and active compensation slows the increase of the pressure build-up and therefore helps to retain the transglottal conditions of phonation for longer. A further compensatory feature is that the [+ voice] consonants are shorter than their [- voice] counterparts as the increase of the supraglottal pressure is lower when duration is shorter. The conditions of voicing during obstruent pronunciation also depend upon further parameters, like the position in the utterance (Westbury & Keating 1986), or the phonetic context (e.g. Stevens 1971; Ohala 1976, 1983).

Previous research has shown universal (summary e.g. by Maddieson 1997), and language-specific (e.g. Shih et al. 1999) characteristics. These may depend on the speakers due to differences in cavity sizes or in the usage of compensatory gestures, etc. Differences were also found in the manners and places of articulation as well (e.g. Jesus & Shadle 2003). The [+ voice]/[- voice] pairs differ from each other in their phonetic voicing based on the type of their phonological opposition, but the phonetic context and the position in the word or in the utterance can decrease this difference or neutralize it (for instance, voicing assimilation, or final devoicing).

Several features have been found to covary with and serve as a secondary cue to the voicing contrast. These are consequences of the timing of glottal and supraglottal gestures (e.g. aspiration, F_1 values at the time of the onset of modal voice, Stevens 1998), or the consequence (adjacent sound duration, for vowels: Slis and Cohen 1969: Dutch; Mair–Shadle 1996: French; Smith 1997: American English; Lousada et al. 2010: European Portuguese) of a parameter characteristic for the contrast (e.g. duration of the consonant: Baum–Blumstein 1987, and Smith 1997: American English; Docherty 1992: British English; Mair–Shadle 1996: French; Jesus–Shadle 2003, and Lousada et al. 2010: European Portuguese). The duration of the preceding vowel is quasi universal, but in some cases it does not hold that the voiced consonant is preceded by longer vowel duration (e.g. Crystal & House 1998; Stevens et al. 1992). The vowel-consonant duration balance (or at least the effort involved in producing it) has been explained by several theories (Chen [1970] proposed an explanation based on muscle movements, Stevens et al. [1992] interpreted it in terms of the timing of supraglottal and glottal gestures).

Various research has been carried out on voicing contrast in Hungarian fricatives, and some on plosives. Other results can also be found which did not primarily aim to analyze the realization of this opposition, but which even so contribute to its description. All these raise further questions on the

[1] Only the phoneme status of /dz/ raises questions (for details see: Siptár 1995).

topic. The difference between plosives and fricatives seems to be obvious (c.f. e.g. Bárkányi–Kiss 2009; Gráczi–Bárkányi 2012 with Gósy–Ringen 2010; Gráczi 2011a) due to the specific feature, i.e. the contradiction between high intensity friction and phonation hypothesis (e.g. Stevens 1998). The effect of the position in the utterance is found to be relevant in both manners of articulation. The vowel context of the fricatives did not affect their phonetic voicing (Bárkányi & Kiss 2009), but it was different from that of the CC-clusters (/v/: Kiss–Bárkányi 2006; /z/: Gráczi 2008). The secondary cues seem to play a greater role/appear more frequently in the case of fricatives, e.g. the duration of the preceding vowel differed in intervocalic context before them, but this was not the case for plosives (compare: Bárkányi & Kiss 2009; Gráczi & Bárkányi 2012 with Gráczi 2011b). However, several authors found differences in the case of plosives, while other research did not show nay difference for either, or in general (cf. Magdics 1966, Kassai 1979 with Olaszy 1994, Kovács 2002, Kohári 2010).[2] Some methodological differences relating to measurement also exist among previous studies, but the results showed similar tendencies. The present thesis aims to find answers to some of the questions raised by previous investigations and the comparison of these studies.

Questions and hypothesis of the thesis

Four experiments are included in this study.

i) The first investigation aims to consider the consequences of the measurement methods. Several questions may arise, but the thesis focused on one, the segment boundary choice, that may affect most data. Our question was to what extent if any it affects the data when three VC- and CV-boundary phenomena are measured as the part of the consonant. The chosen phenomena were the aspiration, preaspiration, and a quasisilent period (that might be voiced or voiceless, but shows the acoustic features of neither the vowel nor the consonant). The "consonant-based" measurement considered the start and end of the closure, or friction, and calculated the phenomena in question as a part of the vowel. The "vowel-based" method calculated the same as a part of the consonant. Several further methods can be found, but these two have appeared in studies of Hungarian (see e.g. Gráczi 2008 vs. Bárkányi & Kiss 2009). All three manners of articulation have been analyzed in the thesis.

a) Our main hypothesis was that the separation of the [+voice] and [-voice] phoneme realizations, based on their voiced part ratio and their duration, will be acceptable regardless of the methodology. We also hypothesized that b) the duration, and c) the voiced part ratio of the consonant would differ significantly when the two measurements were compared. c) A higher ratio of difference was expected in the consonants with frication, due to the intense fricative noise or low airflow during phonation. e) A smaller difference was expected in the duration of the preceding vowel as preaspiration is rare in Hungarian (e.g. Gráczi 2011b, Gráczi & Bárkányi 2012). e) Our assumption was that the differences will be greater in the unvoiced phoneme realizations than in the voiced ones as aspiration and preaspiration can occur in the former.

ii) The effect of the position in the utterance was analyzed for the plosives, and the fricatives. The difference of the effect of pause, and vowel as phonetic context, of the word boundary position was analyzed. a) The voiced part ratio was hypothesized to be higher in the intervocalic word-internal position (in both the voiced and unvoiced phonemes realizations) due to the presence of voicing at the start (unvoiced phonemes), its slower cessations (in both), and the need for voicing restart for the next segment. The lowest ratio of voiced part was expected in utterance final position (preceding a pause, and maybe breath) due to the lack of its restart (Westbury—Keating 1986). b) We assumed that the

[2] Various parameters (like the articulatory configuration of the vowel, as well as the context, the stress position, and the position in the word) were found to affect vowel duration, and most of them were not consistent among the studies (see e.g.: Kassai 1982; Magdics 1966; Olaszy 1994, 1996; Kovács 2002).

difference of the realizations of the voiced counterparts would be smaller in utterance final position, i.e. opposition would be neutralized or weakened more often in these cases based on the voiced part ratio. Results of earlier research (Kiss & Bárkányi 2006; Bárkányi & Kiss 2009; Gráczi & Bárkányi 2012; Gósy & Ringen 2010; Gráczi 2011a) showed that the loss of voicing is greater in utterance final position than in intervocalic context. Based on these data, word internal intervocalic and utterance final positions can be expected to result in different realizations. c) Both the duration of the consonant and the preceding vowel were assumed to contribute to the opposition.

iii) The effect of the phonetic context was analyzed for all three manners of articulation. Three vowels were chosen, and the realizations of the voicing counterparts were analyzed in word internal position in these three contexts. The hypotheses were: a) phonetic voicing, and b) the duration of the consonant were expected to be different across the contexts, and that c) this difference's occurrence would depend on the relative distance between the articulatory configurations of the consonant and the vowel. d) We also hypothesized that the separation of the realizations of the counterparts would be independent of the context, as it affects the voiced and the unvoiced ones in a similar way (if not equally).

iv) Analysis of spontaneous speech was also carried out. The reasons for conducting both laboratory and spontaneous speech analysis were as follows. The frequency of occurrence differs among the phonemes in the latter (for Hungarian data see: Szende 1973; Gósy 2004), therefore the comparison of the counterparts at any place of articulation for all manners of articulation is possible only in some positions. Phenomena that play or may play a role in the realizations are less able to be controlled in spontaneous speech. However, in this speech mode the effect of the frequency of occurrence of both the consonant in question and its counterpart on their realization and on the weakening or loss of opposition could be analyzed. The difference of the speech planning processes of the two speech modes hypothetically result in differences in position on the H&H scale (Lindblom 1990), thus higher variability is expected in spontaneous speech. (The differences in speech style depend on various factors that can be considered to be similar within the analyzed spontaneous speech recordings). We hypothesized that in this speech mode we would also find a) differences in the realizations of the voiced phonemes according to the manner of articulation (higher variability of fricatives was assumed), b) higher variability of realizations in (pause-to-pause) phrase initial and final positions, c) and a higher ratio of devoicing d) and neutralization in utterance final position again. We also hypothesized that the duration of the consonant would be longer in the unvoiced phonemes, following universal tendencies.

Methods of the experiments

To answer the research questions, two types of speech material were used. In one, nonsense words were embedded in carrier sentences in order to tightly control the affecting factors, while in the other speech samples of the BEA database (Gósy 2008) were analyzed. Labeling and analyses were carried out using Praat 5.1—5.3 (Boersma & Weenink 2011), and statistical analysis was performed using SPSS 19.0.

Nonsense word experiments

Speakers, and speech samples

Four women and two men (24–29 ys.) participated in the experiment. All of them speak standard Hungarian.

CVCVC structured nonsense words were embedded in carrier sentences (‘The word [CVCVC] can be seen on the screen.’). One on the consonants was always one of the voicing counterparts in Hungarian, the others were /l/; the vowels were always identical. Five positions were recorded:

- Utterance initial: ||_/ɔ/: C*alal a képernyőn látható szóalak.*
- Intervocalic word initial: /ɔ/#_/ɔ/: *A képernyőn a* C*alal alak látható.*
- Intervocalic word internal: /ɔ/_/ɔ/: *A képernyőn a la*C*al alak látható.*
- Intervocalic word internal: /i/_/i/: *A képernyőn a li*C*il alak látható.*
- Intervocalic word internal: /u/_/u/: *A képernyőn a lu*C*ul alak látható.*
- Intervocalic word final: /ɔ/_#/ɔ/: *A képernyőn a lala*C *alak látható.*
- Utterance final: /ɔ/_||: *A képernyőn látható alak a lala*C.

The plosives (/b, p, d, t, ɟ, c, g, k/), and the fricatives (/v, f, z, s, ʒ, ʃ/) were analyzed in all carrier sentences; the affricates (/d͡z, t͡s, d͡ʒ , t͡ʃ/) only in word internal ones.

The recordings were carried out in a sound-attenuated chamber, with an AT 4040 omnidirectional microphone, and sampled at 44.1 kHz, 16 bit. The sentences were introduced to the participants in a randomized order one-by-one by means of SpeechRecorder (Draxler–Jänsch 2004). Four repetitions per sentence were recorded.

Segmenting methods, analyses

The segment boundaries of the vowels and the consonants, and the ceasation and restart of voicing were labeled manually, based on the oscillogram and spectrogram. The HNR-measurements and the information that appears in the labels were done using Praat scripts. The standard values of Praat for spectrogram visualization were kept, except for the dynamic range (set to 45 dB), and pre-emphasis (set to 3.0 dB/octave). These were changed due to the relative intensity of sibilants, which appear at higher frequencies and result in a less transparent view of vowel formants with the standard values.

The end and the start of voicing were set at the proper voice offset and onset, so breathy (and creaky) voice was also classed as voicing.

The boundaries of the segments were labeled in two ways for the methodological experiment that was carried out only in the /ɔ/_/ɔ/ position. The “consonant-based” measurement considered the start and end of the closure, or friction, and calculated the three previously mentioned phenomena with respect to the vowel. The “vowel-based” method calculated these phanomena with respect to the consonant.

The boundaries for the analyses of the position in the utterance, and of the vowel context, were labeled with the “vowel-based” method.

The following time points were labeled: start of the preceding vowel, start and end of the consonant, end and start of voicing, and in plosives the most intense burst. We also noted if the burst was not apparent or the release was friction, and if the manner of articulation of the realization was not as expected.

The following were calculated/retrieved: i) duration of the vowel, and ii) the consonant, iii) the restart of voicing from the most intense burst (in [-voice] phoneme realizations it is equal to the VOT, but for [+voice] ones we did not calculate this value except if the voicing ceased, therefore we do not call it VOT), iv) the ratio of voiced parts in the consonant, and v) the HNR-values of fricatives. (The mean HNR of the consonant was calculated based on extracting the consonant in question by parabolic windowing, and using the standard settings of Praat.)

Spontaneous speech analysis

Speakers and speech samples

Speech samples of the BEA database (Gósy 2008) were selected. Interviews with 8 female speakers (aged 22—54 years) were analyzed. The recordings were labeled first manually in Praat in a pause-to-pause manner, then they were automatically segmented using MAUS (Shiel et al. 1999). The segmentation was manually corrected. The same positions as in the nonsense speech material were analyzed, but the vowel context was not fixed.

Methods and analyses

The start of the consonant was set to the last zero-crossing before the appearance of the fricative noise, or the start of the closure. The end of the consonant was set to the start of the first vertical striation of the F_2 in the following vowel.

The start and end of the consonant and the voicing was retrieved. The duration of the consonant and its voiced parts ratio were calculated.

Results, and discussion

Comparison of measurement methods

The two different ways of duration labeling showed differences for all plosives and for most of the fricative and affricate consonants (Fig. 1). No such difference was found for the /v/ and /d͡ʒ/ realizations. In the case of the labiodental fricative, the reason is the frequent appearance of sonorant-like speech sounds (HNR-values in the position experiment). The differences were greater and more frequent among fricatives due to the sudden intensity loss after constriction release. All manners of articulation showed greater and more frequent differences in the unvoiced phoneme realizations. The reason is that aspiration appears only in the realizations of these phonemes, while the phase without any higher frequency print can be found in both.

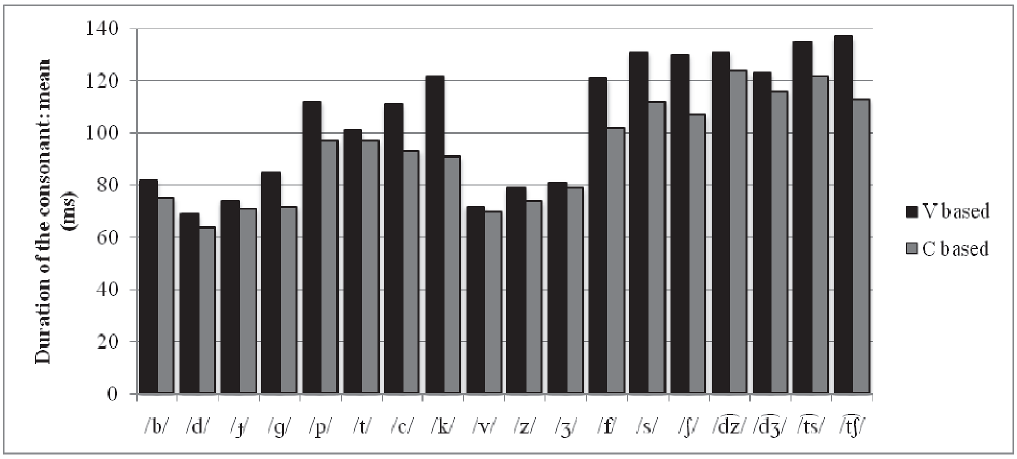


Figure 1: The mean duration of the consonant (ms) with the two measurement methods

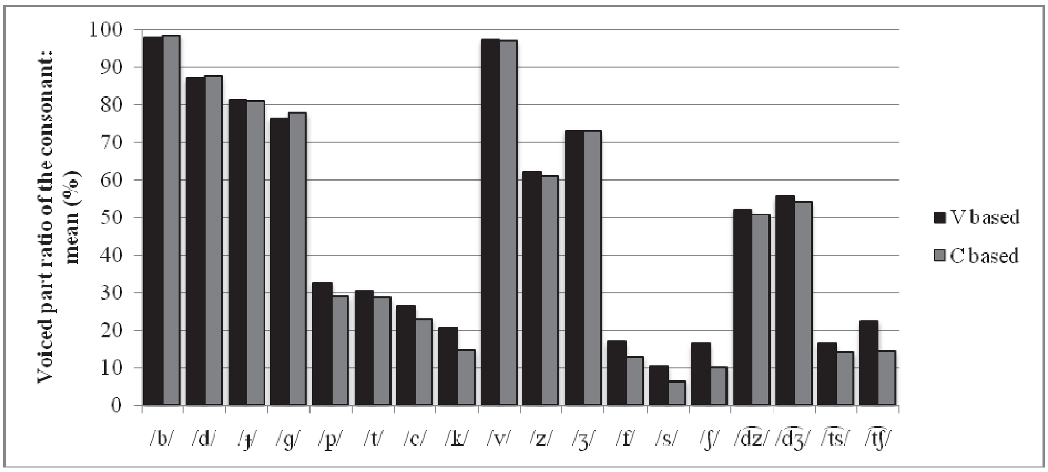


Figure 2: The mean voiced part ratio of the consonant (%) with the two measurement methods

The voiced part ratio (Fig. 2) differences showed diverse results for the three distinct manners of articulation. That of the plosives was different only in the case of the unvoiced bilabial and velar, while for all unvoiced fricatives and for all affricates it was independent of phonological voicing. The smaller difference in the realizations of the voiced phonemes has three causes: the number of the fully voiced cases, the fact that the voicing ceases mostly during the closure or the ‘steady state’ of the friction, and also because the restart of voicing can begin earlier in these cases due to the lower supraglottal pressure.

The duration of the preceding vowel (Fig. 3) appeared to be less dependent upon the measurement method. The reason is that preaspiration is rare, as well as being speaker- and place of articulation-dependent in Hungarian (e.g. Gráczi 2011b; Gráczi & Bárkányi 2012). A difference was found for the preceding vowel of /k/, /f/, and /ʃ/. The dispersion of data is higher for the fricatives calculated with the “vowel-based” method than those calculated with the “consonant-based” one. In affricates only the preceding vowel of /t͡ʃ/ showed a difference in its duration between the two measurement types.

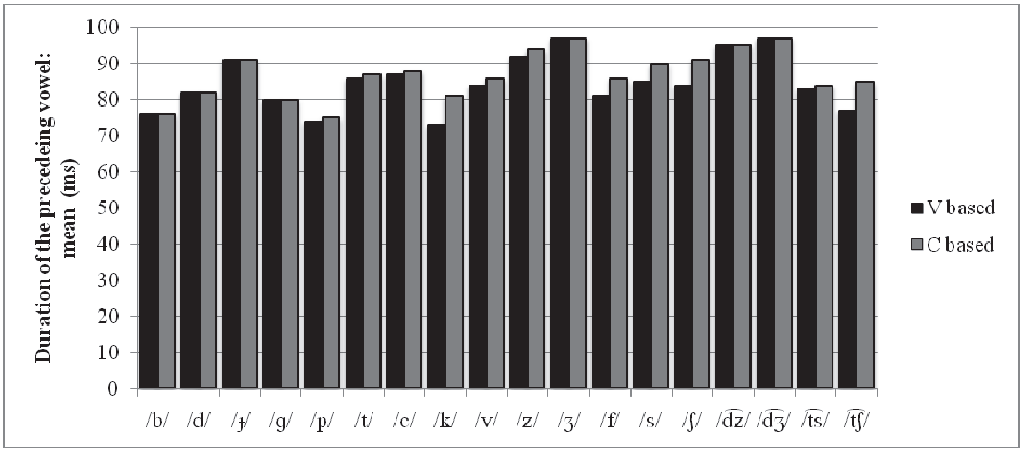
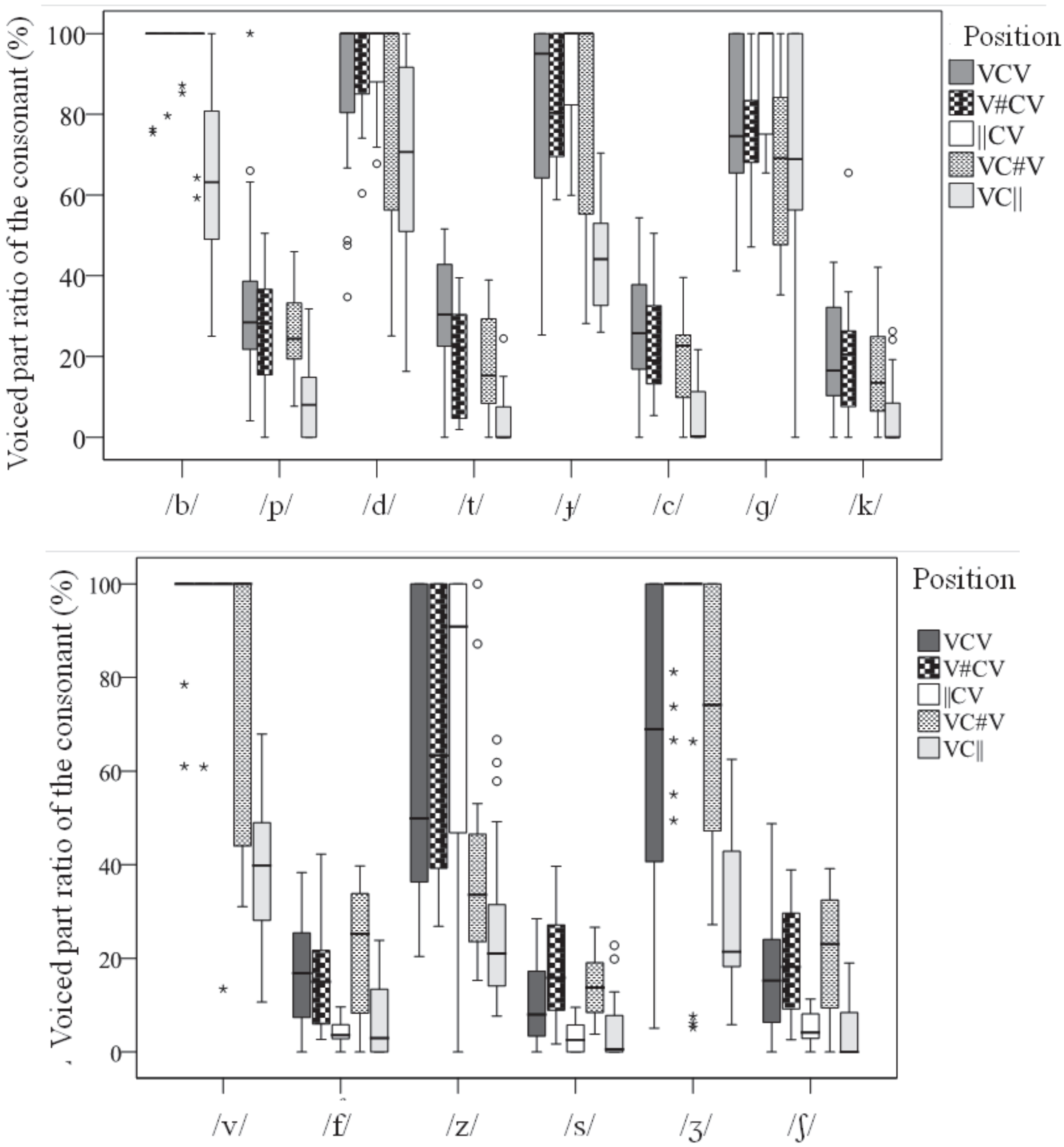


Figure 3: The mean duration of the preceding vowel (ms) with the two measurement methods

Our main hypothesis, that the separation of the counterparts does not depend on the measurement methods, was partially proved. The voiced part ratio of the [+ voice]/[- voice] counterparts was significantly different for both measurement methods. The duration of the plosives and the fricatives followed the universal tendencies in both cases, i.e. the measured duration was shorter for the voiced phoneme realizations. The voiced affricates had been found in earlier studies to vary in their realization; in phonological descriptions the alveolar is considered to be a long sound. The dispersion of both voiced affricates was wide in the present results, therefore though the other two manners of articulation followed the tendency in spite of the high frequency of overlaps, the affricates did not. The ratio of overlap of the voiced part ratio of the members of the counterparts was smaller in the case of the “consonant-based” labeling method, while the duration of fricatives was more stable in the “vowel-based” one.

Position in the utterance

The effect of the position was analyzed in the case of plosives and fricatives. The phonetic context of vowel vs. pause, and that of word-boundary vs. word-internal positions were analyzed. The utterance initial and final positions resulted in a lower voiced part ratio for all voiced consonants compared to the word medial one. The plosives showed similar values in word initial and medial positions, while fricatives tended to have a higher voiced part ratio in the former. The results were dependent upon the place of articulation, the speaker, and the position as well. The plosives with posterior place of articulation tended to devoice to a higher extent than the anteriors, meaning both more frequent occurrence of any length of devoicing and lower voiced part ratios. The voiced labiodental fricative tended to realize more often in a sonorant-like sound (HNR-values), thus it was also more often fully voiced. The postalveolar fricative varied in some positions either with the labiodental, or with the alveolar. In utterance final position all voiced consonants became more (often, and to a lower voiced part ratio) devoiced, and the spirantic feature raised (HNR-values). Therefore the position in the utterance is of great importance—as expected—to the realization. The unvoiced phoneme realizations showed less variablity in the utterance final position, while the interspeaker variability for voiced ones increased when moving from the initial to the final position. Interspeaker variability was apparent in all positions for the fricatives, but for the labiodental one that was dependent upon the speaker only in the word final and utterance final positions.



4. ábra. A mássalhangzó zöngés részének aránya (%) a hangsorbeli pozíció függvényében.

The voiced part ratio showed statistically significant differences in all positions between the members of the [+ voice]/[- voice] pairs, but the frequency of overlap in their realizations increased in word final, and even more in utterance final positions due to the differences in the voiced phoneme realizations described above.

The duration of the consonants followed the universal tendencies in all cases, i.e. the voiced consonants were significantly shorter than their unvoiced counterparts in all positions. The consonant durations showed more frequent overlap between the counterparts in all positions. The absolute duration of the vowel did not conform to expectations (as in some of the earlier studies). In the case of the plosives, only the word and utterance final positions showed a significant difference, while the position and the (consonant) place of articulation played an important role in the case of the fricatives. The ratio of the vowel and the consonant duration (t_V/t_C) was, however, significantly different in all positions for both manners of articulation.

Effect of the vowel context

All three manners of articulation were analyzed regarding the vowel context effect (Table 1). Three vowels were chosen.

The voiced part ratio of the voiced plosives increased as a tendency in the /ɔ/ – /i/ – /u/ order, which was a consequence of decrease in dispersion of the data, while the unvoiced ones showed greater differences only in some cases. The tendency in the voiced ones reached the level of significance only in the case of the velar. None of the other two manners of articulation showed any statistically relevant difference, only tendencies: fricatives showed lower dispersion between /i/-s (but this did not even mean a difference in the mean), and the voiced labiodental showed higher frequency of partially devoiced realizations between /u/-s than in the other two contexts. The affricates also showed dispersion differences only; the tendency is that this is wider in the /ɔ/ context.

Table 1: Results of the realizations with regard of the vowel context (mean and standard deviation)

	Duration of C (ms)			**Voiced part ratio of C (%)**			**Duration of V (ms)**		
	/ɔ/	**/i/**	**/u/**	**/ɔ/**	**/i/**	**/u/**	**/ɔ/**	**/i/**	**/u/**
/b/	97.9± 7.0	95.1± 9.1	94.9±10.5	82±12	79±12	96±17	76±11	67± 7	68±14
/d/	87.1±19.5	95.8±10.2	94.4±10.6	69±12	65±18	71±15	82± 9	75±08	76±15
/ɟ/	81.5±23.3	84.2±21.5	87.7±19.3	74±15	90±21	78±22	91± 9	78±12	81±12
/g/	76.4±18.7	89.1±18.1	87.1±18.2	85±16	88±19	99±24	80±11	73±15	61±11
/p/	32.8±20.9	32.2±16.0	31.1±12.0	112±16	110±16	137±18	74±12	66±11	65±13
/t/	30.3±13.8	29.1±14.0	29.3±16.2	102±15	109±19	111±19	86±13	73±11	79±17
/c/	25.5±17.1	23.8±11.3	31.4±11.9	111±17	130±19	121±19	87±12	71±13	77± 7
/k/	20.8±13.7	19.9± 9.9	24.4±10.4	123±22	123±28	140±21	73±16	69±10	65±12
/v/	97.5± 8.9	98.9±14.2	90.0±19.2	072±10	076±12	100±17	84±18	75±13	79±13
/z/	62.1±29.6	57.5±20.4	74.5±24.7	079±17	091±15	090±24	92±16	89±15	92±15
/ʒ/	69.3±31.5	83.3±23.7	74.2±30.7	081±11	092±17	089±14	97±12	83±16	95±16
/f/	16.8±10.4	21.7± 9.2	23.4± 8.3	121±17	129±17	154±14	81±13	65±08	65±11
/s/	14.5±20.2	16.8± 9.6	17.5± 8.2	131±25	138±15	144±19	85±17	73±13	83±13
/ʃ/	16.3±12.5	15.3± 7.5	18.1± 7.5	132±22	138±11	141±20	82±18	83±16	81±16
/d͡z/	52.2±21.1	49.4±13.5	49.3±16.4	135±25	133±39	137±18	83±12	72±10	79±14
/d͡ʒ/	55.5±23.1	53.5±21.6	61.5±19.8	137±19	139±20	140±30	77±13	70±13	74±14
/t͡s/	16.4±10.3	29.1±27.7	19.0±09.8	131±25	135±36	136±35	95±13	81±14	86±17
/t͡ʃ/	22.3±12.6	16.2±11.3	26.0±17.8	123±30	125±33	119±27	97±12	84±12	82±16

The plosives, excluding the alveolar [+voice]/[- voice] counterparts, showed statistically significant difference in their durations among the three contexts. As expected, the duration changed with the distance between the vowels' and the consonants' articulatory configurations. Bilabials were longer between /u/-s than in the two other contexts, while the palatals were shortest in the /ɔ/ context and longest inbetween /i/-s. The velars were longest between /u/-s of the three contexts. Fricatives were shortest in the /ɔ/ context, while the other two did not show consequent variation (even when the distance of the vowel and the consonant were analyzed). Affricates were the longest between /i/-s.

The duration of the vowel depended—as expected—on its own articulatory configuration. Realizations of /ɔ/ were the longest, while the two high vowels' durations did not show consequent differences before plosives. Difficulties in separating the start of the F_2 in /u/ resulted in high dispersion of its duration values. The realizations of /i/ were clearly shorter than those of /ɔ/ preceding fricatives, but the further variance in the results may depend upon other features. The preceding vowels before affricates showed clearer tendencies. The low vowel was the longest, as already mentioned, and /i/ was realized as the shortest in the case of all affricates. This difference is significant.

The voiced part ratio and the duration of the consonant differed significantly between the voicing pairs in the case of the plosives and the fricatives in all contexts. The voicing counterparts in the affricates also showed different voiced part ratios, regardless of their duration (as the duration variance mentioned above suggested). The voiced consonants overlapped most frequently with their unvoiced counterparts in terms of their voiced parts ratio in the case of the /ɔ/ context. The duration overlaps were very variable (except for /ʒ/ - /ʃ/). The duration of the preceding vowel was not significantly different in any context before the plosive counterparts; the low vowel differed only before the postalveolars in the case of fricatives, while the high ones proved to be longer before all voiced fricatives than unvoiced ones. The vowel duration preceding the affricates was always significantly different with only one exception (/u/ before postalveolars).

Spontaneous speech analysis

The frequency of occurrence of the phonemes does not allow us to prove or reject the hypothesis, but provides tendencies, and augments the results of the nonsense word experiment.

Fricatives underwent devoicing more often and to a higher degree in intervocalic word internal positions in this speech mode (Fig. 5). The effect of the utterance final position was similar in this speech mode, but plosives showed higher dispersion.

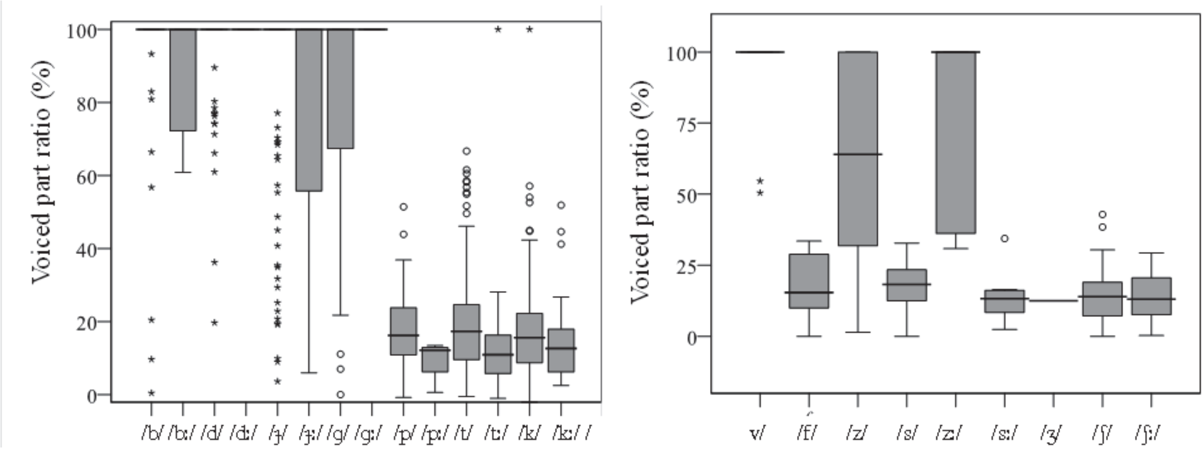


Figure 5: Voiced part ratio of plosives and fricatives in word internal intervocalic position (meadian and dispersion)

The effect of the position in the utterance is apparent in plosives and fricatives (only three voiced affricates appeared). The utterance final position results in lower mean and higher dispersion in voiced part ratio than the word internal one. The plosives varied according to their place of articulation in utterance and word initial position. The voiced part ratio differed according to the place of articulation in all positions for the plosives, but while in word internal position the mean was at least 90% for each one, in utterance initial position it varied between 70% and 100%. The voiced labiodental fricative also showed devoicing in utterance initial position, which resulted a lower voiced part ratio for them than in the nonsense word experiment. Higher frequency of neutralization of the opposition can be expected in utterance final position in this speech mode, e.g. the difference of that in the velar plosives did not reach the level of significance.

We consider it important to mention that the ratio of the voiced part (therefore the weakening or loss of the opposition at the phonetic level) did not seem to vary with the frequency of occurrence of the phoneme or its counterpart. Both velar plosives appear frequently in all positions, the voiced ones' devoicing was frequent in each case, while the voiced postalveolar affricate (two occurrences) showed great distance from its counterpart though a high devoicing ratio (60% of the duration of its realizations was voiced, that is approximately twice that of its counterpart). The voiced postalveolar fricative also occurred infrequently. It showed high variability, which meant that it even appeared in the middle the interval of its counterparts (which were frequent), or it was fully voiced.

The duration of the consonant differs greatly. The shorter duration of the voiced consonants can be accepted as a general tendency, but the difference depends upon several parameters. Taking the velar plosives into consideration again, these overlapped often in terms of their voiced part ratio in word internal position, and showed no overlap in their duration. By contrast, in utterance final position, where the difference of their voiced part ratio did not reach the level of significance, their duration overlapped entirely.

Main differences in the nonsense word and spontaneous speech data

The data for voiced parts ratios showed higher diversity in spontaneous speech at certain places of articulation (in the case of the plosives, at the posterior ones).

Neither the voiced part ratio nor the consonant duration showed similar amount of neutralization in laboratory speech compared to the case of velars in utterance final position in the spontaneous one. The reason might be the difference in the speech styles.

The differences relating to the places and manners of articulation were supported in the present experiments, as well. Both speech modes showed lower (ratio and frequency of) devoicing for plosives, than for fricatives (with the higher variability of the palatal and velar plosives). The combination of the two manners of articulation in fricatives favors the lower ratio of voicing. The voicing ceased rarely in the labiodental fricative in non word, and utterance final positions; it was characteristic for the alveolar, while the results for the postalveolar were diverse. The frequency may be a reason for this, but Bárkányi & Kiss (2009) found that the existence of a minimal pair does not affect the realizations and, as mentioned above, the realizations in the spontaneous speech did not vary with the frequency of their own or their counterpart, but rather with the position in the utterance.

Word frequency was assumed to play a role in several features, like in vowel neutralization, or in the neutralization due to voicing assimilation (summary: Ernestus et al. 2006). The function of the word in the speech/utterance may also play a role in the realizations (e.g. Gósy & Horváth 2010; Dér & Markó 2010; but the results were not different for all analyzed words). The stress patterns also differ in the speech styles[3], and this has also been proved to play a role in phonetic voicing in some languages. The studies analyzed word stress, but the sentence accent cannot be ignored either. These issues were not investigated in the present study, but they may contribute to the differences between the two modes.

Regarding stress patterns, the nonsense words can contribute some relevant data. The V#_V and the -V_V- positions differed in the duration patterns of the preceding /ɔ/ realizations. The preceding vowel differed significantly before the [+ voice]/[- voice] counterparts, though it is itself a definite article and therefore is more likely to be subject to centralization and reduction. The consonant itself is in the stressed syllable in the first case[4].

The frequency of occurrence is also crucial to? the number of analyzable realizations. As mentioned, and expected, in spontaneous speech the phonemes did not appear with the same

[3] Stress and accent can be expressed by several features, and also appear and are perceived in a different way in the two speech modes (Markó 2012), therefore the probability of its effect exceeds the simple question of stress/accent position.

[4] In Hungarian the first syllable is stressed in all non-function words.

frequency, and their number also differed with regard to the position in the utterance. Therefore it is not possible to draw definitive conclusions, but it can be stated that the position in the utterance is of great importance, and that the less frequent consonants tend to follow the more frequent ones, while in some cases the reason for any difference needs further investigation. In the case of the affricates, the combination of the two manners of articulation may contribute to the phonetic voicing of the realizations in addition to the phoneme frequency. This brings us back to the palatal plosives. The unvoiced phoneme did not appear in the spontaneous speech, and the realizations of this phoneme and its voiced counterpart were to some extent different from those of the other plosives (e.g. Kovács 2002; Gráczi 2011a). The reason for this difference is hypothesized to be due to the specific place of articulation (e.g. Ladefoged 1971), and the reason for the diverse realizations of the postalveolars may also be similar.

The difference between the spontaneous and the laboratory speech may arise from the following. In the former, we did not restrict the vowel context (either the vowel itself or the equality of the preceding and the adjacent ones). The nonsense word results showed that the vowel context had the greatest effect on the vowel itself. However, the comparison of the [+voice]/[-voice] counterparts was carried out when they appeared in the same context. The speakers were also different, therefore interspeaker differences might have contributed to a lesser extent to the differences in the results for the two speech modes. The difference of the style of the two speech modes must be also considered, meaning that the speakers may aim to target closer productions in the case of laboratory speech. The spontaneous speech may cooccure more often with fast speech processes (for fast speech processes see Ács & Siptár 1994).

Summary and conclusions

I) The universal tendency in the effect of the place of articulation on the phonetic voicing was proved, but the resulting differences depend also on the manner of articulation. The plosives undergo devoicing less frequently and to a lower ratio than fricatives, as expected, as in the former only the closure results in pressure build up that may lead to the cessation of voicing, while in the latter the intense friction and the low airflow during phonation are also contradictory features. The posterior place of articulation leads to a lower voicing ratio in plosives, while /ʒ/ in the case of fricatives shows diverse realizations.

II) The effect of the position in the utterance partially depends on the consonant itself (fricatives in utterance and word initial positions), but the word final and utterance final positions lead to a higher ratio of devoicing in both the plosives and fricatives, this phenomenon is enhanced in the latter case due to its aerodynamic features and articulatory settings (Westbury & Keating 1986).

III) The effect of the phonetic context on the consonant voicing and duration mainly depends on the distance between the articulatory configurations of the vowel and the consonant.

a) The context affects phonetic voicing to a lower ratio, and almost does not influence the separation of the [+voice]/[-voice] counterparts based on this parameter.

b) The context effect on the consonant duration is greater. This may be as a consequence of the timing of the articulatory gestures and/or the distance between the configurations. The universal tendency for the duration of the [+voice]/[-voice] counterparts to differ in the same context is not influenced by these differences.

IV) The duration of the preceding vowel and its varying with the phonological voicing of the consonant depend mainly on its own articulatory configuration. It is important to note that the analyzed low vowel has no phonological length counterpart that would differ only in one gesture. Therefore a greater freedom could be hypothesized in its realization than in the other two analyzed vowels, which both have length counterparts differing only in one gesture. A reason might be the closer tongue height of the vowels to the consonants, resulting in a difference in the intergestural timing. In addition, the tempo of pressure build up might contribute to the different duration patterns of the vowels.

V) The absolute duration of the vowels did not in all cases follow the expected pattern (see above), but the ratio of this and the duration of the consonant were different at a significant level in almost all cases.

a) The absolute duration of the vowel showed higher variability in relation to the phonological voicing of the consonant, reaching statistical significance in all cases in utterance final position. The duration of the unvoiced consonants was longer than that of the voiced ones in all positions, therefore articulatory timing cannot be the only reason for the vowel duration patterning. In the utterance final position, the voiced part ratios of the consonants showed higher frequency of overlap. A reason for this might be an attempt to retain the opposition. Also, final lengthening might be enhanced in the case of longer durations.

VI) Only minor differences can be found between the results of the spontaneous speech and nonsense word reading results. The main reason for these differences might be the different speech styles, but several other phenomena might also play a part. This hypothesis can be proved or rejected only with further experiments.

VII) Neither the frequency of the phoneme, nor its counterparts seem to play an important role in phonetic voicing. Earlier investigations found that the existence of minimal pair does not affect this characteristic either (e.g. Bárkányi & Kiss 2009), while the lexical frequency plays a role in the degree of neutralization caused by voicing assimilation (e.g. Ernestus et al. 2006).

Reference

Ács Péter – Siptár Péter 1994. Túl a gondozott beszéden. In Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 2. Fonológia*. Budapest: Akadémiai Kiadó. 550–578.

Bárkányi Zsuzsanna – Kiss Zoltán 2009. Word-final fricative contrasts in Hungarian. A phonetic approach. Előadás a *BuPhoC* 2009 nov. 5-i ülésén. Letölthető: http://budling.nytud.hu/~cash/papers/ buphoc09-slide.pdf

Baum, S. R. – Blumstein, S. E. 1987. Preliminary observations on the use of duration as a cue to syllable-initial fricative consonant voicing in English. *Journal of the Acoustical Society of America* 82: 1073–1077.

Boersma, P. – Weenink, D. 2011. *Praat: doing phonetics by computer* (Version 5.1.13). Elérhető: http://www.fon.hum.uva.nl/praat/download_win.html

Chen, M. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22: 129–159.

Crystal, T. H. – House, A. S. 1988. Segmental durations in connected-speech signals: Current results. *Journal of Acoustic Society of America* 83: 1553–1573.

Dér Csilla Ilona – Markó Alexandra 2010. A pilot study of Hungarian discourse markers. Language and Speech 50(2): 135–180.

Docherty, G. J. 1992. *The Timing of Voicing in British English Obstruents.* Berlin: Foris Publications..

Draxler, C. – Jänsch, K. 2004. SpeechRecorder, a universal platform independent multi-channel audio recording software. In *Proceedings of LREC*, Lisbon. 559–562.

Ernestus, M. – Lahey, M. – Verhees R., F. – Baayen, H. 2006. Lexical frequency and voice assimilation. *Journal of the Acoustical Society of America* 120(2): 1040–1051.

Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Budapest: Osiris.

Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszédkutatás 2008*. 194–207.

Gósy, Mária – Horváth, Viktória 2010. Changes in articulation accompanying functional changes in word usage. *Journal of the International Phonetic Association* 40(2): 135–163.

Gósy, Mária – Ringen, Catherine O. 2009. Everything you always wanted to know about VOT in Hungarian. Elhangzott: *International Conference on the Structure of Hungarian* 2009, Budapest, 2009. szeptember 1. Elérhető: http://icsh9. unideb.hu/pph/handout/Ringen_Gosy_handout.pdf.

Gráczi Tekla Etelka 2005. Suttogott mássalhangzók zöngésségének észlelése. Beszédkutatás 2005. 78–90.

Gráczi Tekla Etelka 2008. Alveoláris spiránsok akusztikai fonetikai vizsgálata. *Beszédkutatás 2008.* 33-51.

Gráczi Tekla Etelka 2011a. Explozívák a zöngésségi oppozíció függvényében. Váradi Tamás (szerk.): *V. Alkalmazott Nyelvészeti Doktoranduszkonferencia*. Budapest: MTA Nyelvtudományi Intézet. 51–66.

Gráczi Tekla Etelka 2011b. Voicing contrast of intervocalic plosives in Hungarian. In Lee, W.-S. – Zee, E. (szerk.): 759–762.

Gráczi Tekla Etelka – Bárkányi Zsuzsanna 2012. Voiced fricatives in Hungarian in utterance-final position. In *Proceedings of Grammar and Context – New Approaches to the Uralic Languages III.* (megjelenés alatt).

Jesus, L. M. T. – Shadle, Chr. H. 2003. Temporal and devoicing analysis of European Portuguese fricatives. In Solé, M.-J. – Recasens, D. – Romero, J. (szerk.): 779–782.

Johnson, K. 2012. *Acousic and Auditory Phonetics*. Malden, MA: Wiley–Blackwell.

Joos, M. 1968. The isolation of styles. In Fishman, J. A. (szerk.): *Readings in the Sociology of Language.* The Hague: Mouton. 185–191.

Kassai Ilona 1979. Időtartam és kvantitás a magyar nyelvben. *Nyelvtudományi Értekezések* 112. Budapest: Akadémiai Kiadó.

Kassai Ilona 1982. A magyar beszédhangok időtartamviszonyai. In Bolla Kálmán (szerk.): *Fejezetek a magyar leíró hangtanból.* Budapest: Akadémiai Kiadó.

Kiss Zoltán – Bárkányi Zsuzsanna 2006. A phonetically-based approach to the phonology of [v] in Hungarian. *Acta Linguistica Hungarica* 53(2): 175–226.

Kluender, K. R. – Diehl, R. L. – Killeen, P. R. 1987. Japanese Quail can learn phonetic categories. *Science* 237: 1195–1197.

Kluender, K. R. – Lotto, A. J. 1994. Effects of first formant onset frequency on [-voice] judgments result from general auditory processes not specific to humans. *Journal of the Acoustical Society of America* 95: 1044–1052.

Kohári Anna 2010. *Magyar magánhangzók kontextusfüggő időviszonyai*. Szakdolgozat, ELTE BTK.

Kovács Magdolna 2002. *Tendenciák és szabályszerűségek a magánhangzó-időtartamok produkciójában és percepciójában.* Debrecen: Debreceni Egyetem Kossuth Egyetemi Kiadója.

Ladefoged, P. 1971. *Prelimineries to linguistic phonetics.* Chicago & London: The University of Chicago Press.

Lindblom, B. 1990. Explaining phonetic variation: a sketch of the H&H theory. In. Hardcastle, W. J. – Marchal, A. (szerk.) 1990. 403-440.

Lousada, M. – Jesus, L. M. T. – Hall, Andreia 2010. Temporal acoustic correlates of the voicing contrast in European Portuguese stops. *Journal of the International Phonetic Association* 40(3): 261–276.

Maddieson, I. 1997. Phonetical universals. In Hardcastle, William J. – Laver, John (szerk.): 619–639.

Magdics Klára 1966. A magyar beszédhangok időtartama. *Nyelvtudományi Közlemények* 68: 125–139.

Mair, S. J. – Shadle, C. H. 1996. The voiced/voiceless distinction in fricatives: EPG, acoustic and aerodynamic data. *Proceedings of Institute of Acoustics* 18: 163–169.

Markó Alexandra 2012. A magyar hangsúly realizációinak és észlelésének összefüggése felolvasásban és spontán beszédben. In Markó Alexandra (szerk.):. 2012. 277–303.

Ohala, J. J. 1976. A model of Speech Aerodinamics. *Report of the phonological Laboratory*. Berkeley.

Ohala, J. J. 1983. The origin of the sound patterns in vocal tract constraints. In MacNeilage, P. F. (szerk.): The Production of Speech. New York: Springer Verlag. 189–216.

Ohala, J. J. 1997. Aerodynamics of phonology. In *Proceedings 4th Seoul International Conference on Linguistics*, Seoul. 92−97.

Olaszy Gábor 1994. Sound duration measurements in declarative sentences. *Acta Linguisica Hungarica* 42: 51–62.

Shadle, C. H. 1990. Articulatory acoustic relationships in fricative consonants. In Hardcastle, W. J. – Marchal, A. (szerk.): cím 1990. 187–209.

Shadle, C. H. 1991. The effect of geometry on source mechanisms of fricative consonants. *Journal of Phonetics* 19: 409–424.

Shadle, C. H. 1997. The aerodynamics of speech. In Hardcastle, William J. – Laver, John (szerk.) cím 33–64.

Shevoroshkin, V. V. 1971. O dvukh [v] v russkom iazyke. In Vysotskii, S. S. – Panov, M. V. –Reformatskii, A. A. – Sidorov, V. N. (szerk.) *Razvitie fonetiki sovremennogo russkogoazyka*. Moscow: Nauka. 279-286.

Shiel, F. – Draxler, C. – Hoole, P. – Tillmann, H. G. 1999. New Resources at BAS: Acoustic, Multimodal, Linguistic. *Proceedings of Eurospeech*. 2271–2274.

Shih, C. – Möbius, B. – Narasimhan, B. 1999. Contextual effects on consonant voicing profiles: A cross-linguistic study. *Proceedings of the 14th International Congress of Phonetic Sciences* (San Francisco, CA) 2: 989–992.

Siptár Péter 1995. A magyar mássalhangzók fonológiája. *Linguistica, Series A, Studia et Dissertationes* 18. Budapest: MTA Nyelvtudományi Intézet.

Slis, I. H. – Cohen, A. 1969. On the complex regulating the voiced-voiceless distinction I. *Language and Speech* 12: 80–102.

Smith, C. L. 1997. The devoicing of /z/ in American English: Effects of local and prosodic context. *Journal of Phonetics* 25(4): 471–500.

Stevens, K. N. 1971. Airflow and turbulence noise for fricative and stop consonants. Static considerations. *Journal of the Acoustical Society of America* 50: 1180–1192.

Stevens, K. N. 1998. *Acoustic Phonetics*. Cambridge, Massachusetts–London: The MIT Press.

Stevens, K. N. – Blumstein, S. E – Glicksman, L. – Burton, M. – Kurowski, K. 1992. Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. *Journal of Acoustic Society of America* 91(5): 2979–3000.

Szende Tamás 1973. Spontán beszédanyag gyakorisági mutatói. *Nyelvtudományi Értekezések* 81. Budapest: Akadémiai Kiadó.

Westbury, J. R. – Keating, P. A. 1986. On the naturalness of stop consonant voicing. *Journal of Linguistics* 22: 145–166.

My publications on realization of voicing (contrast)

In journals:

Gráczi Tekla Etelka 2005. Suttogott mássalhangzók észlelése. *Beszédkutatás 2005.* 78–95.

Gráczi Tekla Etelka 2008. Alveoláris spiránsok akusztikai fonetikai vizsgálata. *Beszédkutatás 2008.* 33–52.

Markó Alexandra – Gráczi Tekla Etelka – Bóna Judit 2009. Zöngésségi hasonulás a spontán beszédben és a felolvasásban (esettanulmányok). Beszédkutatás 2009. 5–27.

Gráczi Tekla Etelka 2010. A spiránsok zöngésségi oppozíciójának néhány jellemzője. *Beszédkutatás 2010.* 42–56.

Markó Alexandra – Gráczi Tekla Etelka – Bóna Judit: The realization of voicing assimilation rules in Hungarian spontaneous and read speech: Case studies. *Acta Linguistica Hungarica* 57/2–3. 210–238.

Gráczi Tekla Etelka 2011. Intervokális explozívák a zöngésségi oppozíció függvényében. *Beszédkutatás 2011.* 46–60.

Gráczi Tekla Etelka 2012. Obstruenskapcsolatok a spontán beszédben a zöngésségi oppozíció függvényében. *Beszédkutatás 2012.* 30–42.

In proceedings:

Gráczi Tekla Etelka 2008. „Z"-öngétlenedés: Beszélőfüggő paraméter? In Váradi, Tamás (szerk.): *II. Alkalmazott Nyelvészeti Doktorandusz Konferencia.* Budapest: MTA Nyelvtudományi Intézet. 34–44.

Bóna, Judit – Gráczi, Tekla Etelka – Markó, Alexandra 2008. Coarticulation rules and speaking style dependency. Rudolph Sock, Suse, Yvess Laprie (szerk.): *Proceedings of International Seminar for Speech Production 2008.* Strasbourg, 245–248.

Bóna Judit – Gráczi Tekla Etelka – Markó Alexandra 2009.: A koartikulációs szabályok érvényesülése a különböző beszédstílusokban. In Gecső Tamás – Sárdi Csilla (szerk.): *A kommunikáció nyelvészeti aspektusai. Segédkönyvek a nyelvészet tanulmányozásához 99.* Székesfehérvár, Budapest: Kodolányi János Főiskola, Tinta Könyvkiadó. 50–55.

Gráczi Tekla Etelka 2011. Explozívák a zöngésségi oppozíció függvényében. *V. Alkalmazott Nyelvészeti Doktorandusz Konferencia.* 51–66.

Gráczi Tekla Etelka 2011. Voicing contrast of intervocalic plosives in Hungarian. *Proceedings of 17th International Congress of Phonetic Sciences.* 759–762.

In press

Gráczi Tekla Etelka –Bárkányi Zsuzsanna 2012. Voiced fricatives in Hungarian in utterance-final position. *Grammar and Context – New Approaches to the Uralic Languages III.* (in press).

Gráczi Tekla Etelka 2012. Spiránsok zöngéssége spontán beszédben. Váradi Tamás (szerk.): *VI. Alkalmazott Nyelvészeti Doktoranduszkonferencia.* Budapest: MTA Nyelvtudományi Intézet. (in press)

My other publications on obstruents

Gráczi Tekla Etelka 2007. Az explozívák realizációja mássalhangzó-kapcsolatok első tagjaként. *Beszédkutatás 2007.* 35–47.

Gráczi Tekla Etelka – Kohári Anna 2012. A zöngekezdési idő egy módszertani kérdés függvényében. Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngekezdési időig.* Budapest: ELTE Bölcsészettudományi Kar és MTA Nyelvtudományi Intézet. 228–248.

Number of all publications: 36 studies (2 book chapters), and 7 (conference and book) reports