Inter-speaker and intra-speaker variability indicating a synchronous speech sound change

Mária Gósy

Introduction

Analysis of fine acoustic-phonetic details of speech has confirmed the variability of speech production across various speakers and within the same speaker on various occasions. Speaker-characteristic differences are supposed to be rooted in long-term voice features based on anatomical characteristics, and in short-term voice features based on the individual articulatory gestures of speaker (Hollien 1990, Johnson 1997, Mooshammer et al. 2008). Speech sound articulation changes in time, both in one's life and across generations (Gósy 1999). The fact that speakers articulate speech sounds or words differently, depending on dialect, phonetic context, prosody, speech style, speaking tempo, gender, hearing acuity, sociolinguistic factors, emotions or even on smoking habits, was demonstrated and discussed during the past decades (eg Heđever et al. 1999, Pierrehumbert 2001, Shiller et al. 2002, Gelfer & Mikos 2005, Recasens & Espinosa 2006, Lindblom et al. 2009, Brunner et al. 2011, Varga 2012). Such variability seems to be crucial in automatic speech recognition and in forensic speaker identification (eg McDougall 2006, Kahn et al. 2011).

Decades ago, Lindblom emphasized that physiological and biomechanical factors have a decisive role in the motor control of articulation (1983). According to his theory, speech output varies along a hyper- and hypospeech continuum. In addition to all these factors, many others might play an important role in the variability of speech production, such as learning or social environmental factors (Perkell et al. 1997). Both interspeaker and intra-speaker variability is an important topic in phonetics,

both in speech production and speech perception (eg Johnson 1997, Dankovičová & Nolan 1995, Mooshammer et al. 2008).

The pronunciation of speech sounds in spontaneous speech and their perception seems to be paradoxical when we consider both inter- and intra-speaker variability. Speakers always try to approach an assumed target speech sound, but they are influenced by so many factors that they may or may not reach the intended target. In everyday communication listeners constantly map the acoustic-phonetic patterns of all actually produced speech sounds to the supposed targets, an exercise that requires a flexible, modifiable "neural spectrogram" stored in the brain (see the results of testing topographical mapping of vowels on the cortical surface: Obleser et al. 2003).

Ohala pointed out that pronunciation variations are responsible for sound change (1993, 2012) and claims that instrumental phonetic studies shed light on the large amount of variations existing in speech sound articulation. His assumption is that speech sound changes are initiated from the listeners because they might perceive some speech sounds erroneously and so they interpret the speaker's intended articulation gesture differently from his/her intended one. In the seventies, historical sound changes were explained by the speakers' effort to attain articulation economy (eg Benkő 1988, Lindblom 1990, Ohala 2012). Wardhaugh's view is that the two most important factors in linguistic change are the existing variations and the pressure of society as well as their interrelations (1995). Phonetic changes are assumed to occur in words rather than in speech segments (see Pierrehumbert 2001, Bybee 2008). Speech sound changes seem to be gradual in time; however, their rapidity may be different based on various factors influencing them in a certain era (eg Kiss & Pusztai 2005). It is difficult – if not impossible – to define the actual stage of a synchronous change on a time scale. However, we can characterize certain parameters of a speech sound in order to compare them to those measured in the past and in this way it becomes possible to evaluate the extent and direction of change. This requires some reliable parameters from the past that can be compared with the same parameters as they appear at present (see Pettigrew 1990).

The Hungarian speech sound inventory contains a vowel in whose two articulation gestures seem to be undergoing a change at the present time. Traditionally, phoneticians used to characterize the vowel *á*[1] as a back, unrounded, downmost (lowest), long vowel (eg Vértes O. 1982, Bolla

---

[1] We will use the letter symbol for identification of the vowel *á* in this paper, since some of its phonetic features seem to be uncertain, and so it was not obvious at the beginning of this study which of the IPA-symbols would be appropriate to use.

1995, Kassai 1998, Gósy 2004). In addition, some authors, from Wolfgang von Kempelen (1989 [1791]) on, have claimed that the vowel *á* is not articulated as back in the oral cavity as other Hungarian back vowels are (eg Bolla 1995, Kovács 2004, Gósy 2004, Gráczi & Horváth 2010, Beke & Gráczi 2010). Instrumental measurements in the recent past provide some support for the claim that in articulating this vowel the tongue occupies a more front position. If so, it must also be asked whether the parameters of this vowel show large individual differences indicating uncertainty in its articulation, which might lead to changes in articulatory gestures.

Our research questions are (i) whether the inter-speaker and intra-speaker differences in the pronunciation of the *á* vowel are really so large as to support the claim that the articulatory gestures are changing, and (ii) whether the inter-speaker differences exceed intra-speaker differences in the articulation of this vowel. If we can confirm that acoustic variability occurring both among different speakers and within the same speaker in the articulation of the *á,* this could be regarded as the beginning of a speech sound change. Our hypothesis is that there is a synchronous change taking place in the articulation of the vowel *á,* affecting the horizontal position of the tongue in the oral cavity. We suppose that there will be at least three groups of vowels (both among females and males) indicating that there is a tendency away from traditional back articulation towards front articulation. If we can categorize the measured data accordingly, we may prove the existence of a synchronous change in the articulation gestures of the vowel *á.*

Subjects, method, material

Spontaneous speech samples were used from the (Hungarian) BEA Speech database (Gósy 2012). Subjects spoke about their work, hobbies and told their opinions about some current issues (the educational system, taxes, mobile phones and children, smoking habits, new driving rules in the country and so on). The duration of the recorded speech samples varied across speakers depending on the required number of *á* vowels we could use in the analysis (the mean duration per speaker was 26 minutes). 20 to 40 vowels were analyzed in the narratives of 14 females and 14 males (ages between 22 and 28), altogether 614 realizations in the females' speech samples and 695 realizations in the males' speech samples.

The first three formants of the vowel *á* were measured in the first and second syllables of the words (both in monosyllables and in polysyllables). The following criteria were taken into consideration when selecting the vowels: (i) the vowel should be perceived without doubt as a Hun-

garian *á* vowel like in the word *lát* '(s)he sees', which means that other vowel quality realizations, for example the neutral vowels, were excluded from the analysis, and (ii) the vowels should occur both in content and in function words (without any further selection). We did not intend to (and could not) control either the syllable type or the number of syllables in the words. The vowel quality of the *á* vowels was defined by the author and another phonetician (in a few cases of disagreement, the vowels in question were excluded). Examples of words that contained this vowel: *fák* 'trees', *már* 'already', *látogatókkal* 'with visitors', *támad* '(s)he attacks', *bármelyik* 'whichever'; *órákban* 'in hours', *kutyám* 'my dog', *inkább* 'rather', *találkoztunk* 'we met', *egymáshoz* 'to each other'.

Measurements of the formants were carried out manually in the middle of the steady-state phase of the vowel considering the visual information of both the spectrograms and oscillograms as well as the repeated audition of the vowel in question (using Praat software: Boersma & Weenink 2011). In some cases we also used the automatic formant curves generated by the software. In addition, the energy spectra of the vowels were also used (FFT-analysis, Fast Fourier Transformation) in all vowels to support the values of the three formants.

Statistical analysis was carried out by SPSS 17 software (one-way ANOVA, Tukey post-hoc tests). Fuzzy clustering technique was used for automatic classification.

Results

Vowels are characterized acoustically by their formant structures. Traditionally, the first three formants are considered out of which the first two formants (F1 and F2) alone specify the vowel quality in many languages while third formants seem to be more closely linked to the speakers' individual characteristics (eg Ladefoged 1975, Hollien 1990, Künzel 1995). Formant measurements of the Hungarian vowels go back to the sixties of the past century (Tarnóczy 1965, Magdics 1965). During the past decades several values were defined for the first three formants of the vowels, using mainly read speech (Molnár 1970, Vértes O. 1982, Olaszy 1985, Bolla 1995, etc). These studies provided means and sometimes ranges on the formant values of the *á* representing frequently only a few (sometimes one or two) speakers' data.

The first formant of the vowel indicates the tongue position vertically in the oral cavity while the second formant indicates the tongue position horizontally in the oral cavity. The higher value of F1 shows lower tongue position while the higher value of F2 shows that the tongue is placed some-

where in the front of the oral cavity (eg Fant 1973, Slifka 2005). The mean values of the first formant of *á* were about 950 Hz in the case of females and about 780 Hz in the case of males while the second formants ranged about 1500 Hz in the case of female speakers and about 1400 Hz in the case of male speakers in the papers published in the last third of the past century (Magdics 1965, Tarnóczy 1965, Molnár 1970, Vértes O. 1982, Olaszy 1985). These formant values indicate that about half a century ago Hungarian *á* was a vowel articulated at the back of the oral cavity both in the case of females and males and that its tongue position was somewhat lower than that experienced with Hungarian low vowels.

For our data, statistical analysis was carried out to learn whether there is any difference in the formant values of the female speakers between the monosyllables and polysyllables on the one hand, and between the first and second syllables of the polysyllabic words, on the other. No significant differences were found for the first two formants. Although the third formant values differed significantly depending on their occurrence in the first or the second syllable (one-way ANOVA: $F(1, 613) = 36,805$; $p = 0,01$); however, the mean difference is only 111 Hz (the difference of standard deviation is 18 Hz) which is irrelevant for the speakers' sound identification. Therefore, all values were used in analyzing the inter-speaker and intra-speaker variability for the vowel *á* (table 1). Looking at the range of the values, we can postulate large individual differences for all formants. The first formant values around 500 Hz mark a vowel where the tongue height is somewhat lower than in the case of the Hungarian mid-vowels like [o] while higher values indicate that the vowels' tongue position is similar to those vowels that are traditionally low vowels. The F1-values around 900 Hz, however, refer to a vowel where the tongue height is very low, much lower than in the case of the traditionally low Hungarian vowels. The values of F2 around 1600 Hz represent vowels where the tongue is located in the middle of the oral cavity. On the contrary, values around and above 2000 Hz refer unquestionably to a front vowel.

table 1: The first three formant values of the vowel á *in females' articulation*

Formants	Female speakers' values (Hz)			
	mean	std. dev.*	min.	max.
F1	766	81	500	998
F2	1841	96	1600	2200
F3	2788	232	2020	3849

* standard deviation

Similarly to females' data, statistical analysis was carried out concerning the possible effects of monosyllabic and polysyllabic words, as well as the first and the second syllable position of the polysyllabic words in the case of male speakers. Again, no significant differences were found for the first two formants. The third formant values showed significant differences depending on the first and second syllables (one-way ANOVA: $F(1, 694) = 28,036; p = 0,01$); however, the differences here were even lower than those in females (mean difference was 74 Hz while the difference of standard deviation is 15 Hz). Therefore, again all values were used in further analyzing the inter-speaker and intra-speaker variability for the *á* vowel (table 2).

table 2: The first three formant values of the vowel á *in males' articulation*

| Formants	Male speakers' values (Hz)			
	mean	std. dev.*	min.	max.
F1	687	68	488	973
F2	1481	114	1180	1773 **(1936)
F3	2384	183	1406	2958

* standard deviation
** the value in brackets is an outlier

The male speakers articulated their *á* vowels similarly to female speakers considering the tongue height parameter. Their first formants range widely indicating that the tongue height varies from almost mid-tongue position to a very low position. As opposed to the female speakers' second formant values, males' F2-s are much lower than those of the females'. The values, around 1200 Hz, indicate a back vowel in terms of the horizontal position of the tongue in the oral cavity while those between 1500 Hz and 1700 Hz are characteristic of a central vowel where the tongue is positioned in the middle of the oral cavity. With the exception of the neutral vowel, which is not a member of the phoneme inventory of the language, this is not a usual articulation gesture in Hungarian.

Figure 1 shows the female speakers' data for the first formants. The different F1-values indicate the different heights of the tongue. With some speakers, eg 1, 2, 6 and 13, the first formants are lower than with other speakers, eg 3, 4, 8, 10, 14, indicating a lower tongue height with the former and higher tongue height with the latter. Statistical analysis confirmed significant differences among speakers (one-way ANOVA: $F(13,651) = 373,521$;

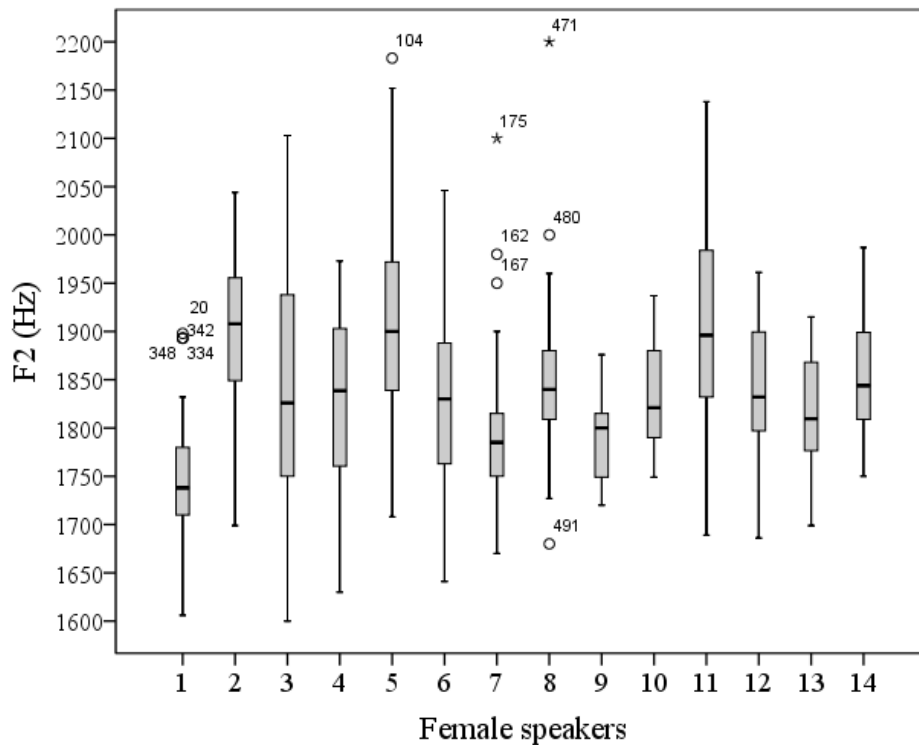*figure 1: Female speakers' first formant values of the vowel á*
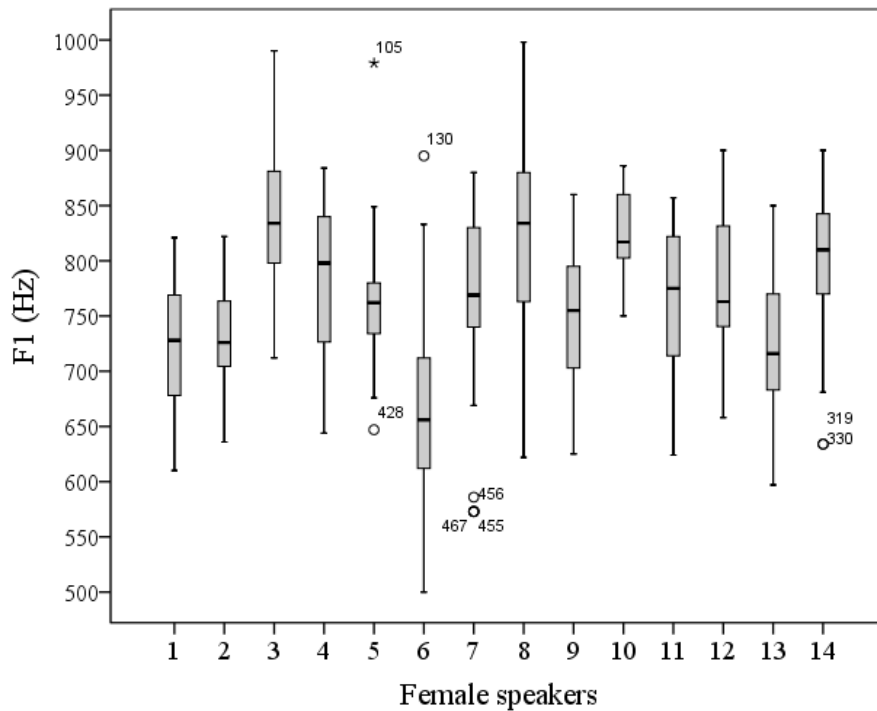
$p = 0.001$). The detailed data of the Tukey post-hoc tests show that our female speakers differ from each other in various ways.

Speakers 1, 3, 6, 8, 10 and 13 differ significantly from 7 or more other speakers, while speakers 2, 4, 5, 7, 9, 11, 12 and 14 differ from 4 or more other speakers. There are speakers whose first formant values do not differ significantly from a group of other speakers, like those of speaker 1 from 5, 7 and 9, or those of speaker 10 from those of 3, 4, 8 or 14. The boxplots of the figure 1 demonstrate also the large intra-speaker variability. There are speakers that seem to be more consistent in their articulation gestures than others (eg speakers 2, 5 or 10 as opposed to speakers 4, 8 or 11). For example, the range of F1 in speaker 2 is about 60 Hz for most of her *á* vowels, while the same range is about 200 Hz in speaker 8.

The statistical analyses also revealed some significant differences in the second formants of female speakers (one-way ANOVA: $F_{(13, 651)} = 1297,202; p = 0.001$). The Tukey post-hoc tests showed as many statistically relevant inter-speaker differences for the F2-values as in the case of the F1-values (figure 2). The second formant values shed light on the fact that our

female speakers articulate the vowel *á* as a palatal one – with the exception of three speakers (1, 7 and 9) – where the F2 values appear under 1800 Hz. The majority of the second formants of the three exceptional speakers are lower than 1800 Hz; however, there are only a few vowels whose F2s appear at those frequencies that indicate mid-position of the tongue in the oral cavity.

figure 2: Female speakers' second formant values of the vowel á

Detailed analysis of the male speakers' data shows similarities to and differences from the females' data. The first formant values demonstrate significant inter-speaker differences (one-way ANOVA: $F(13, 694) = 9,718$; $p = 0.001$). The Post-hoc Tukey tests, however, did not show as large inter-speaker variability as experienced with the females. Some speakers (2, 3, 5, 10 and 12) differ significantly only from one or two other speakers (figure 3).

Four male speakers were found whose first formant values were significantly different from those of six or more other subjects. The first for-

*figure 3: Male speakers' first formant values of the vowel* á
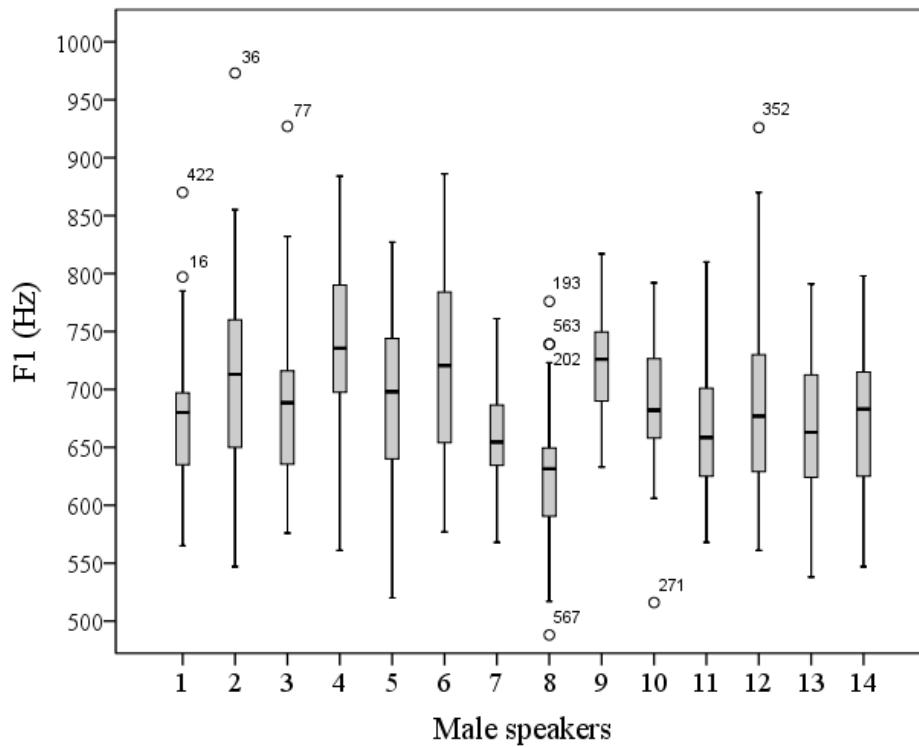
mant values indicate that tongue height in the majority (71.4%) of the male speakers is in some degree higher than traditionally reported for this vowel (see, eg Bolla 1978, Olaszy 1985). Thus, the inter-speaker variability of F1-values suggests that tongue moves in a very wide area in the oral cavity.

Figure 4 shows the male speakers' second formant values where the statistical analysis revealed again significant differences depending on speakers (one-way ANOVA: $F(13, 694) = 21,972$; $p = 0.001$). The Tukey post-hoc test confirmed more significant differences among speakers than we experienced it in the case of the first formants.

All male speakers differed from at least 4 other speakers, while 9 speakers differed from 5 or more other participants. Since the second formants indicate the tongue position in the oral cavity, we can see that only a few data fall into that frequency range which traditionally corresponds to back Hungarian vowels. There are three subjects (1, 8, 13) who can be distinguished from all the others depending on their relatively low F2-values, and another three subjects (6, 9, 14) who show higher F2-values than all the others. Eight speakers' data fall between the data of the six speakers

*figure 4: Male speakers' second formant values of the vowel* á
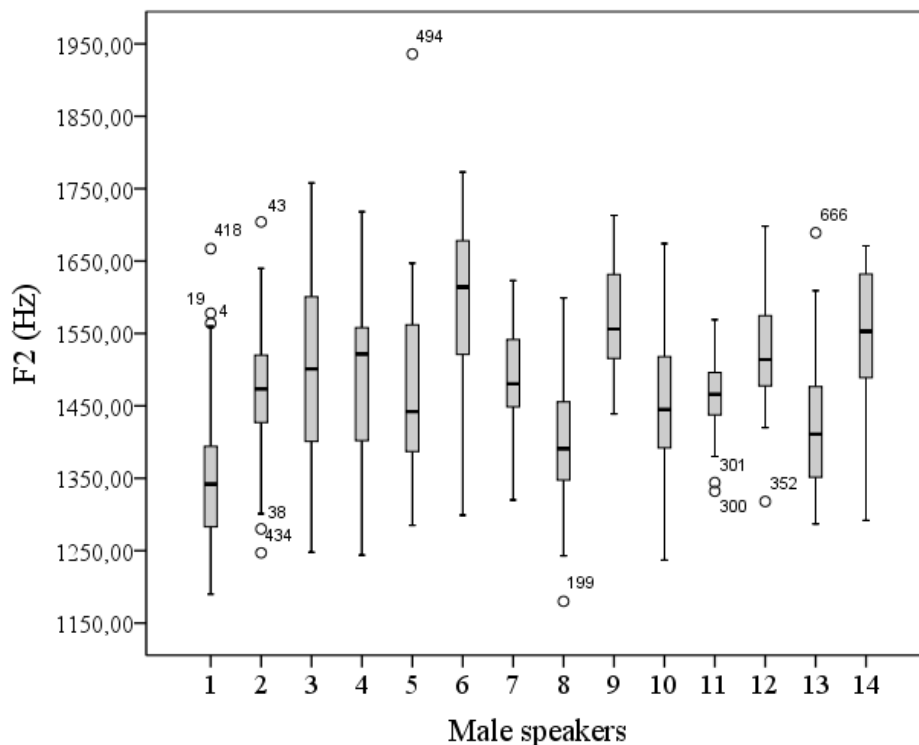
mentioned above, and they show considerable differences supported by the significant results of the post-hoc tests (eg 10 and 12).

We may draw the conclusion that male speakers pronounce the vowel *á* variably, between the back and the mid areas of the oral cavity, and this is more pronounced in the mid cavity. The same applies if we are looking at the intra-speaker data. Three male speakers (6, 9, 14) seem to clearly prefer the mid cavity, while one speaker (1) seems to articulate the vowel *á* as a real back vowel. The other speakers' tongue placement is mainly in the mid region of the oral cavity, but some of their data suggest backward articulation.

Considering all females' first formant values, the ranges show large intra-speaker differences with the frequency ranges of about 130 Hz and about 360 Hz (figure 5). The differences of the second formants are similar to those of F1 in terms of the speakers' articulation consistency. In some speakers (3, 4, 5, 6, 11) tongue placement may range in a relatively large domain in the oral cavity, while in others tongue placement occurs within a relatively narrow space (1, 9, 10 or 13). There are F2-values occurring

within a frequency range of about 450 Hz for the former type of female speakers and about 200 Hz for the latter type. Considering all the data from each speaker, intra-speaker differences are shown to be spectacular (figure 5).

figure 5: *The minimum and maximum values of the first two formants in female speakers*
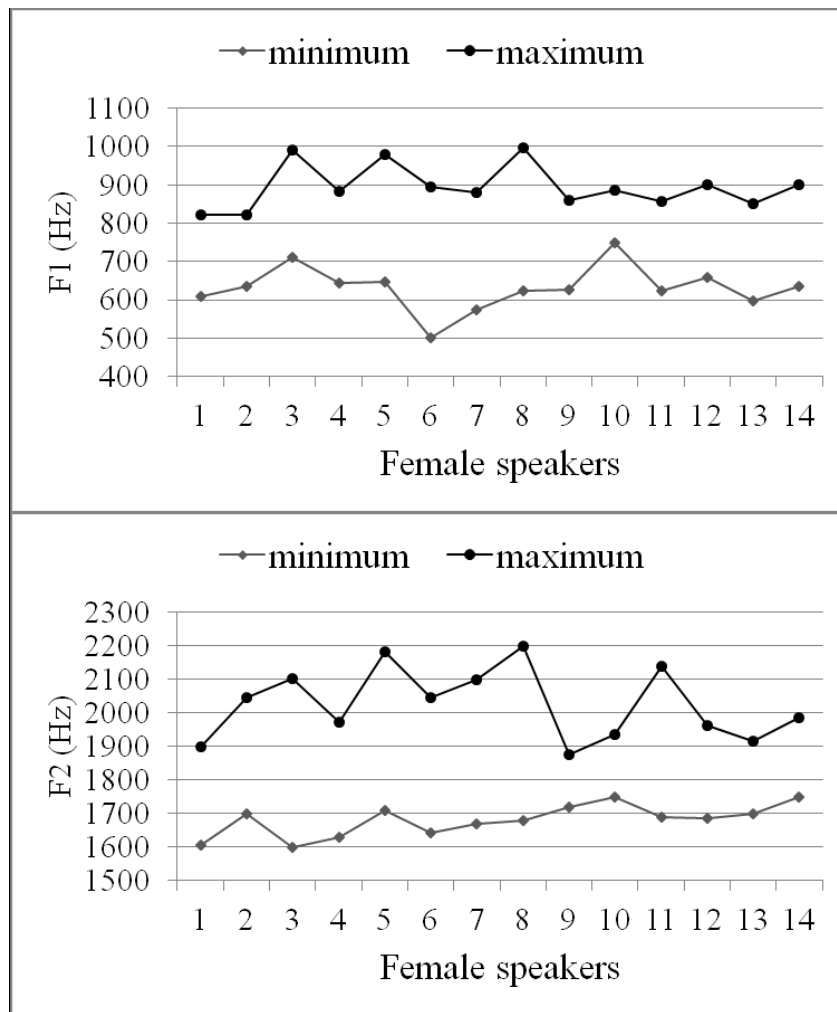
The within-speaker variability in males appears to be high. Most values occur within a frequency range of about 400 Hz for the less consistent male speakers while within a frequency range of about 250 Hz for the more consistent ones in the case of the first formant values (figure 6). Again, there are less consistent male speakers (with a frequency range of about 550 Hz) and more consistent male speakers (with a frequency range of about 300 Hz) when analyzing their second formant values, eg 5 as opposed to 11 (see figure 6).

figure 6: *The minimum and maximum values of the first two formants in male speakers*
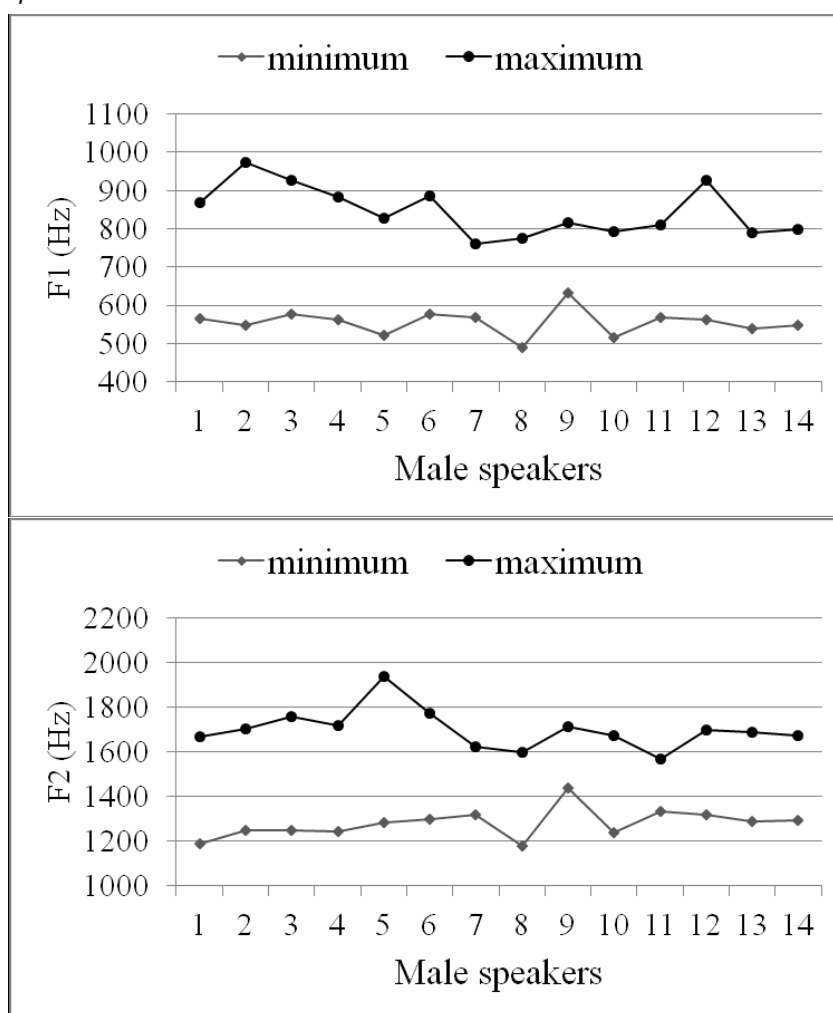
Third formants are claimed to be relatively stable across vowel tokens within a single speaker (Monahan & Idsardi 2010). Although in Hungarian it is also the first two formants that define vowel quality, we analyzed the third formants to see whether they show inter-speaker and intra-speaker variability. The females' third formants (see figure 7) vary in a relatively wide frequency range, mainly between 2500 Hz and 3100 Hz, the difference is statistically confirmed among speakers (one-way ANOVA: $F(13, 613) = 7,507$; $p = 0.001$). The Tukey post-hoc tests, however, did not yield significant differences in about two-third of the cases.

figure 7: Female speakers' third formant values of the vowel á
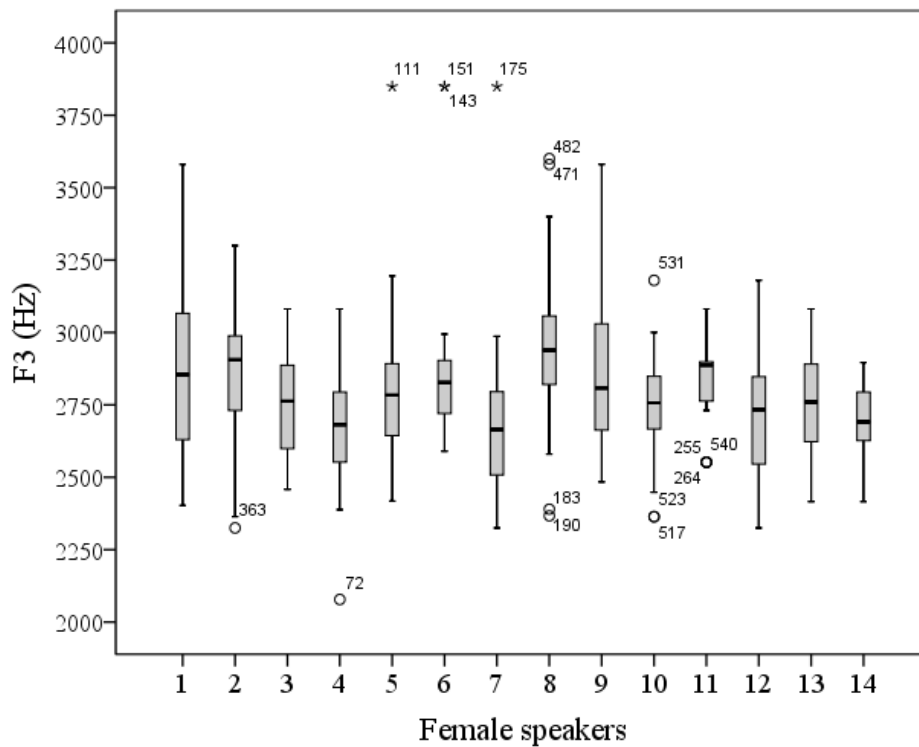
The seemingly large inter-speaker differences go together with similarly large intra-speaker differences resulting in considerable overlaps between speakers. There are only two females demonstrating relatively narrow frequency ranges for their third formants (about 400 Hz) while the F3-values of the majority of our female speakers show a frequency range

*figure 8: Male speakers' third formant values of the vowel* á
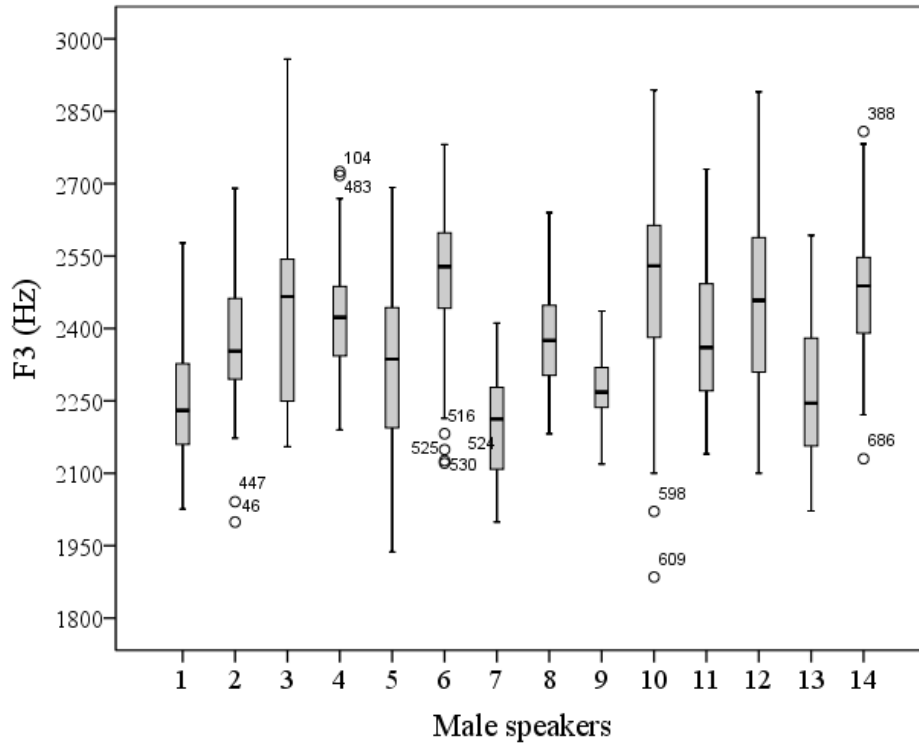
of about 800 Hz. Similar ratios of third formant values were found with the male speakers (see figure 8).

The values show enormous differences among the male speakers supported also by statistical analysis (one-way ANOVA: $F(13, 694) = 19{,}634$; $p = 0.001$). The frequency range differences are larger here (about 1000 Hz) than those found with the female speakers. The Tukey post-hoc tests confirmed larger individual differences among male speakers than in the case of female speakers. Intra-speaker differences are also large, resulting in overlaps among the male speakers as well. Our data show that third formants seem to be more characteristic of the male than of female speakers; however, our findings did not confirm the relatively stable frequency values for these formants (cf Monahan & Idsardi 2010).

Analyzing the speakers' first and second formant values, subgroups of vowels emerged both among females and males. We used the fuzzy clustering technique which is an unsupervised method used to partition the data elements into homogenous subgroups based on similarity distance measure. The elements do not necessarily belong to just one cluster

in this technique. The fuzzy C-means (FCM) algorithm (see Bezdek 1981) was used in this study to cluster the vowels into four subgroups where each one represents a bunch of typical realizations of the vowel *á*. The elements (which are the first two formant values in our case) in the FCM were normalized between 0 and 1. In the FCM algorithm each element has a degree that represents the similarity of one or more clusters. The degree of similarity increases towards the center of the subgroup (marked by the normalized values of the formants), see figure 9. These results identified four groups of vowels both in females and males based on the F1- and F2-values of the vowels. The subgroups and the circles indicate the large variability of vowels.

figure 9: Clustering results of vowels articulated by females (above) and by males (below) demonstrating four subgroups based on their normalized formant values
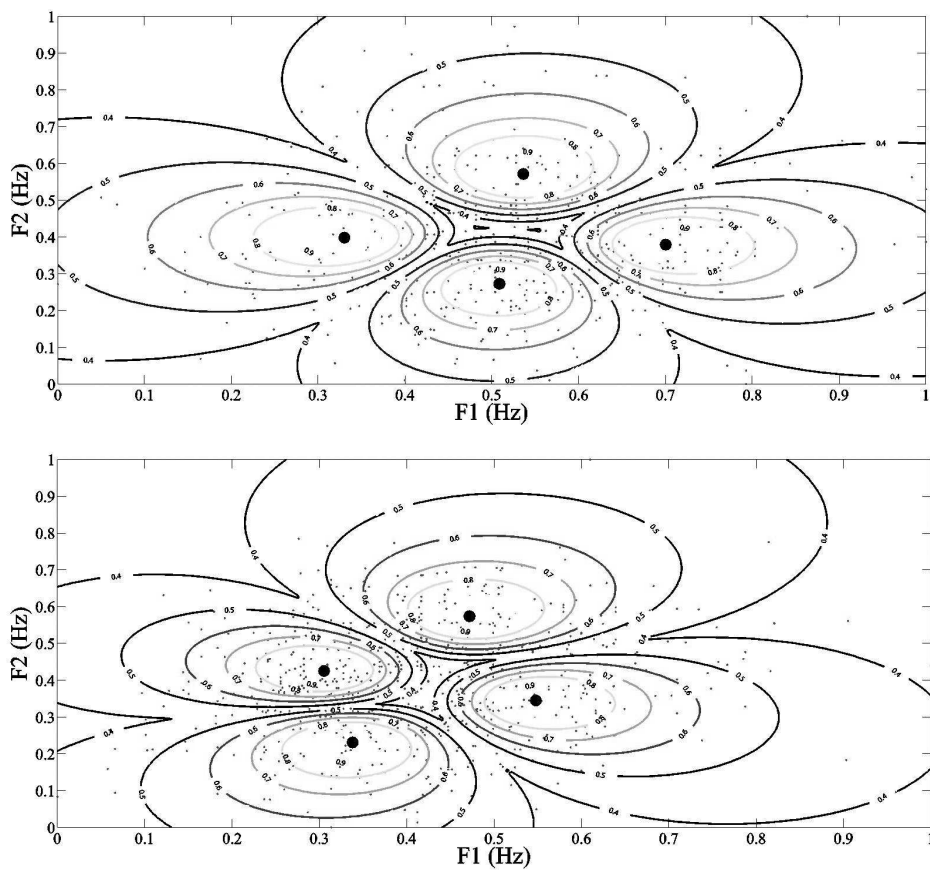
Conclusions

We analyzed the formant structure of the Hungarian vowel *á* based on spontaneous speech samples of 28 young adults. Our research focused on the inter-speaker and intra-speaker variability in the articulation of this vowel in terms of the acoustic consequences of horizontal and vertical tongue movements. The data obtained show that this traditionally back vowel is articulated in the mid of the oral cavity by males while in the front of the oral cavity by females. The values of the first two formants support the claim that, phonetically, it is not a back vowel. Although there are various (methodological) difficulties in comparing our data to those published in the literature decades ago, there can be no question that a synchronous change in the articulation of this vowel is in progress. Our claim is further supported by acoustic-phonetic data from old female speakers who seem to prefer back articulation of the vowel *á* (Bóna 2009). A change in horizontal tongue movement towards the front oral cavity in the articulation of a vowel was observed in 1972 in the Detroit area by Labov and colleagues. A similar sound change was reported about a Turkish vowel whose articulation moved from the back of the oral cavity to the front (Kiliç & Öğüt 2004).

The finding that inter-speaker differences show some gender dependency is worth further discussing. Differences between the formant values of the same speech sound between females and males can be interpreted as a consequence of their anatomical differences, the supposed "open" articulation of female speakers or their faster reaction to the newly invented articulation patterns (see Henton 1995, Diehl et al. 1996). Although several differences can be found in verbal communication between women and men (see, for instance, Eichhoff-Cyrus 1997, Lakoff 2004), no differences have been found in their verbal ability (Hyde & Linn 1988). There are several possible explanations for the seemingly slower articulation change in male speakers as opposed to more robust change in female speakers. Some researchers think that males follow the "articulation norm" more than females do. There is evidence that women take a leading role in sound change (Eckert 1989, Labov 1994); however, Trudgill found the opposite in most sound changes in Norwich (1972). Obviously, gender itself should not have an effect on sound change, but sometimes the gender of speakers who influence a sound change might be important. Based on our data, we can state that young Hungarian female speakers articulate the vowel *á* as a front vowel more than do males. Thus, this study supported the leading role of women in a sound change; however, further research is needed before making universal claims.

Since both groups of females and males in our study showed changes in their articulation gestures, we seek some explanation for this fact. The theory of perceptually triggered sound change (Blevins 2006, Ohala 2012) in our case cannot be a proper explanation since the vowel *á* is the only one articulated with the lowest tongue height, so there is no possibility to misperceive or to confuse it with another vowel. In addition, the change in horizontal tongue movement in this case would not be perceived by the listeners for the same reason. The large intra-speaker differences found support the gradual modification of the neural spectrograms from the old gestures toward new ones. Although at present less faith is placed in explanations of the sound change in terms of linguistic economy, we think that in this case such an explanation cannot be *ab ovo* excluded, at least as one of the factors influencing the change.

We assumed that a change is going on in the articulation of the vowel *á.* In addition, we supposed that there would be groups of vowels demonstrating different phases of this change. Analyses of the first two formants confirmed our hypotheses, which provide support for the existence of a synchronous change in the articulation gestures of this vowel. As expected, great variability was found among speakers, however, more importantly, intra-speaker variability was shown to be essentially as great as the inter-speaker variability. The final question is whether both the inter-speaker and intra-speaker variability is due to the synchronous sound change or this is simply a sign of the great variety of articulation gestures used by speakers in approaching a target vowel. Further analyses of the acoustic consequences of pronunciation are needed to find answers to this question.

REFERENCES

Beke András and Gráczi Tekla Etelka. 2010. A magánhangzók semlegesedése a spontán beszédben. In: Navracsics Judit (szerk.), *Nyelv, beszéd, írás. Pszicholingvisztikai tanulmányok I.* Veszprém: Pannon Egyetem. 57–64.

Benkő Loránd. 1988. *A történeti nyelvtudomány alapjai.* Budapest: Akadémiai Kiadó.

Bezdek, James C. 1981. *Pattern recognition with fuzzy objective function algorithms.* New York: Plenum Press.

Blevins, Juliette. 2006. A theoretical synopsis of Evolutional Phonology. *Theoretical Linguistics* 32 : 117–166.

Boersma, Paul and David Weenink. 2011. *Praat: doing phonetics by computer.* Retrieved on 25 October 2011 from www.fon.hum.uva.nl/praat/download_win.html.

Bolla Kálmán. 1978. A magyar beszédhangok akusztikai analízise és szintézise. *Magyar Fonetikai Füzetek* 1 : 53–67.

Bolla Kálmán. 1995. *Magyar fonetikai atlasz. A szegmentális hangszerkezet elemei.* Budapest: Nemzeti Tankönyvkiadó.

Bóna Judit. 2009. Az idős életkor tükröződése a magánhangzók ejtésében. *Beszédkutatás 2009*: 76–87.

Brunner, Jana, Satrajit Ghosh, Philip Hoole, Melanie Matthies, Mark Tiede, and Joseph Perkell. 2011. The influence of auditory acuity on acoustic variability and the use of motor equivalence during adaptation to a perturbation. *Journal of Speech, Language, and Hearing Research* 54 : 727–739.

Bybee, Joan. 2008. Formal universals as emergent phenomena: the origins of structure preservation. In: Jeff Good (ed.), *Linguistic universals and language change.* Oxford: Oxford University Press. 108–121.

Dankovičová, Jana and Francis Nolan. 1999. Some acoustic effects of speaking style on utterances for automatic speaker verification. *Journal of the International Phonetic Association* 29 : 115–229.

Diehl, Randy L., Björn Lindblom, Kathryn A. Hoemeke, and Richard P. Fahey. 1996. On explaining certain male-female differences in the phonetic realization of vowel categories. *Journal of Phonetics* 24 : 187–208.

Eckert, Penelope. 1989. The whole woman: Sex and gender differences in variation. *Language Variation and Change* 1 : 245–267.

Eichhoff-Cyrus, Karin (ed). 1997. *Adam, Eva und die Sprache.* Mannheim: Dudenverlag.

Fant, Gunnar. 1973. *Speech sounds and features.* Cambridge, MA & London: The MIT Press.

Gelfer, Marylou Pausewang and Victoria A. Mikos. 2005. The relative contributions of speaking fundamental frequency and formant frequencies to gender identification based on isolated vowels. *Journal of Voice* 19 : 544–554.

Gósy Mária. 1999. Az egyéni hangszínezet és a beszélő felismerésének kísérleti-fonetikai megközelítése. *Magyar Nyelvőr* 123 : 424–438.

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

Gósy, Mária. 2012. BEA — A multifunctional Hungarian spoken language database. *The Phonetician* 105 / 106 : 50–61.

Gráczi Tekla Etelka and Horváth Viktória. 2010. A magánhangzók realizációja spontán beszédben. *Beszédkutatás 2010*: 5–16.

Heđever, Mladen, Gordana Kovačič, and Višnja Barišić. 1999. Utjecaj pušenja i radnog staža na osnovni laringalni ton nastavnica. *Govor* XVI: 33–45.

Henton, Caroline. 1995. Cross-language variation in the vowels of female and male speakers. In: *Proceedings of the XIIIth International Congress of Phonetic Sciences.* Stockholm: Stockholm University. 420–423.

Hollien, Harry. 1990. *The acoustics of crime.* New York, London: Plenum Press.

Hyde, Janet S. and Marcia C. Linn. 1988. Gender differences in verbal ability: a meta-analysis. *Psychological Bulletin* 104 : 53–69.

Johnson, Keith. 1997. Speech perception without speaker normalization: An exemplar model. In: Keith Johnson and John W. Mullenix (eds.), *Talker variability in speech processing.* San Diego: Academic Press. 145–166.

Kahn, Juliette, Nicolas Audibert, Jean-François Bonastre, and Solange Rossato. 2011. Inter- and intra-speaker variability in French: an analysis of oral vowels and its implication for automatic speaker verification. In: *Proceedings of the XVIIth International Congress of Phonetic Sciences.* Hong Kong: University of Hong Kong. 1002–1005.

Kassai Ilona. 1998. *Fonetika.* Budapest: Nemzeti Tankönyvkiadó.

Kempelen Farkas. 1989 [1791]. Az emberi beszéd mechanizmusa, valamint a szerző beszélő-gépének leírása. Ford. Mollay Károly. Budapest: Szépirodalmi Könyvkiadó.

Kiliç, Mehmet Akif and Fatih Öğüt. 2004. A high unrounded vowel in Turkish: is it a central or back vowel? *Speech Communication* 43 : 143–154.

Kiss Jenő and Pusztai Ferenc (szerk.). 2005. *Magyar nyelvtörténet.* Budapest: Osiris Kiadó.

Kovács Magdolna. 2004. Pros and cos about Hungarian [a:]. *Grazer Linguistische Studien* 62 : 65–75.

Künzel, Herman J. 1995. Field procedures in forensic speaker recognition. In: Jack Windsor Lewis (ed.), *Studies in general and English phonetics. Essays in Honour of Professor John D. O'Connor.* London: Routledge. 68–85.

Labov, William. 1994. *Principles of linguistic change I. Internal factors.* Oxford: Basil Blackwell.

Labov, William, Malcah Yaeger, and Richard Steiner. 1972. A quantitative study of sound change in progress. Philadelphia: U.S. Regional Survey.

Ladefoged, Peter. 1975. *A course in phonetics.* New York: Harcourt Brace Jovanovich.

Lakoff, Robin. 2004. Language and a woman's place. In: Mary Bucholtz (ed.), *Language and a woman's place.* New York: Oxford Universtity Press. 39–76.

Lindblom, Björn. 1983. Economy of speech gestures. In: Peter MacNeilage (ed.), *The production of speech.* New York: Springer. 217–245.

Lindblom, Björn. 1990. Explaining phonetic variation: a sketch of the H and H theory. In: William J. Hardcastle and Alain Marchal (eds.), *Speech production and speech modeling.* Dordrecht: Kluwer. 403–440.

Lindblom, Björn, Randy Diehl, and Carl Creeger. 2009. Do 'Dominant Frequencies' explain the listener's response to formant and spectrum shape variations? *Speech Communication* 51 : 622–629.

Magdics Klára. 1965. *A magyar beszédhangok akusztikai szerkezete.* Nyelvtudományi Értekezések 49. Budapest: Akadémiai Kiadó.

McDougall, Kirsty. 2006. Dynamic features of speech and characterization of speakers: towards a new approach using formant frequencies. *Speech Language and the Law* 13 : 89–126.

Molnár József. 1970. *A magyar beszédhangok atlasza.* Budapest: Tankönyvkiadó.

Monahan, Philip J. and William J. Idsardi. 2010. Auditory sensitivity to formant ratios: Toward an account of vowel normalization. *Language and Cognitive Processes* 25 : 808–839.

Mooshammer, Christine, Pascal Perrier, and Susanne Fuchs. 2008. Speaker-specific patterns of token-to-token variability. *Journal of the Acoustical Society of America* 123 : 3076.

Obleser, Jonas, Thomas Elbert, Aditi Lahiri, and Carsten Eulitz. 2003. Cortical representation of vowels reflects acoustic dissimilarity determined by formant frequencies. *Cognitive Brain Research* 15 : 207–213.

Ohala, John. 1993. The phonetics of sound change. In: Charles Jones (ed.), *Historical linguistics: Problems and perspectives.* Harlow: Longman. 237–278.

Ohala, John J. 2012. The listener as a source of sound change (perception, production, and social factors). In Maria-Josep Solé and Daniel Recasens (eds.), *The initiation of sound change.* Amsterdam: John Benjamins. 21–36.

Olaszy Gábor. 1985. *A magyar beszéd leggyakoribb hangsorépítő elemeinek szerkezete és szintézise.* Nyelvtudományi Értekezések 121. Budapest: Akadémiai Kiadó.

Perkell, Joseph S., Melanie L. Matthies, Harlan Lane, Frank H. Guenther, Reiner Wilhelms-Tricarico, Jane Wozniak, and Peter Guiod. 1997. Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication* 22 : 227–250.

Pettigrew, Andrew M. 1990. Longitudinal field research on change: Theory and practice. *Organization Science* 1 : 267–292.

Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition, and contrast. In Joan Bybee and Paul Hopper (eds.), *Frequency effects and the emergence of linguistic structure.* Amsterdam: John Benjamins. 137–157.

Recasens, Daniel and Aina Espinosa. 2006. Dispersion and variability of Catalan vowels. *Speech Communication* 48 : 645–666.

Shiller, Douglas M., Rafael Laboissière, and David J. Ostry. 2002. Relationship between jaw stiffness and kinematic variability in speech. *Journal of Neurophysiology* 88 : 2329–2340.

Slifka, Janet. 2005. Acoustic cues to vowel-schwa sequences for high front vowels. *Journal of the Acoustical Society of America* 118 : 2037.

Tarnóczy, Tamás. 1965. Acoustic analysis of Hungarian vowels. *Quarterly Progress and Status Report* 1 (Speech Transmission Laboratory–KHT, Stockholm): 8–12.

Trudgill, Peter. 1972. Sex, covert prestige, and linguistic change in the urban British English of Norwich. *Language in Society* 1 : 179–195.

Varga László. 2012. Van-e magyar mellékhangsúly? In: Markó Alexandra (szerk.), *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngekezdési időig.* Budapest: ELTE, MTA Nyelvtudományi Intézet. 35–49.

Vértes O. András. 1982. Az artikuláció akusztikus vetülete. In: Bolla Kálmán (szerk.), *Fejezetek a magyar leíró hangtanból.* Budapest: Akadémiai Kiadó. 155–165.

Wardhaugh, Ronald. 1995. *Szociolingvisztika.* Budapest: Osiris–Századvég.

Mária Gósy
gosy.maria@nytud.mta.hu
Research Institute for Linguistics,
Hungarian Academy of Sciences