

## ARTICULATORY STUDIES IN HUNGARY – PAST, PRESENT AND FUTURE

ALEXANDRA MARKÓ

ELTE Eötvös Loránd University; MTA–ELTE Lendület Lingual Articulation Research Group  
marko.alexandra@btk.elte.hu  
<https://orcid.org/0000-0003-0301-7134>

TAMÁS GÁBOR CSAPÓ

Budapest University of Technology and Economics; MTA–ELTE Lendület Lingual Articulation Research Group  
csapot@tmit.bme.hu  
<https://orcid.org/0000-0003-4375-7524>

MÁRTON BARTÓK

MTA–ELTE Lendület Lingual Articulation Research Group  
bartokmarton@gmail.com

TEKLA ETELKA GRÁCZI

Research Institute for Linguistics; MTA–ELTE Lendület Lingual Articulation Research Group  
graczi.tekla.etelka@nytud.hu

ANDREA DEME

ELTE Eötvös Loránd University; MTA–ELTE Lendület Lingual Articulation Research Group  
deme.andrea@btk.elte.hu  
<https://orcid.org/0000-0001-5395-8309>

### Abstract

Articulatory studies performed in Hungary date back to the sixties, when different methods were applied for the description of the segment inventory of Hungarian and various other languages (e.g. Russian, German, English, Polish). Palato- and linguography, labiography, and X-ray were used in the analyses of both typical and atypical speech. However, coarticulation, which requires dynamic methods, was not analysed until recently, when the suitable tools and methods, electromagnetic articulography, ultrasound tongue imaging and electroglottography became also available in Hungary. The paper presents an overview of the main issues of articulatory studies on Hungarian in the past and the present. It summarizes the main findings from some studies on gemination and degemination, transparent vowels, phonatory characteristics of emotion, and gives a couple of examples of possible and future applications.

**Keywords:** articulatory studies, applications, Hungarian, overview

### 1. Introduction

Speech researchers have been studying articulation (the co-ordinated movements of the speech organs) and the acoustic features of speech since the 1700s (Kempelen 1791/1989). Special equipment is needed to examine the movements of the articulators (vocal folds, tongue, lips) as most of these structures cannot be seen during the production of speech. As the

study of articulation requires complex instrumentation and all articulation-focused measurement methods modify speech production to some extent, the question arises as to why articulation should be examined at all. As opposed to articulation, the acoustic signal is readily observable and analysable, and the acoustic analysis of speech does not introduce any major measurement-related artefacts to the data. Furthermore, the tools of acoustic analysis are easily available and are relatively inexpensive. For the necessity of articulatory research, the Quantal Theory provides an explanation (Stevens 1989). According to the Quantal Theory of speech production and perception, articulatory gestures have acoustic consequences; however, the magnitude of articulator displacement is not proportional to the magnitude of change observable in the acoustic signal, as the relationship between articulation and its acoustic consequences is not linear. It follows that a minor change in articulation might result in a significant change in the acoustic output if it takes place in the so-called critical region and vice versa, major differences in articulation might not induce any differences in the acoustic signal. The adjective *quantal* refers to this non-linear relationship (Stevens 1989). The importance of the study of articulation is further highlighted by the fact that although certain acoustic parameters highly correlate with articulation, the acoustic parameters cannot be identified with specific articulatory movements in a one-to-one manner. It is commonly known that although the first formant (the first resonance) of the vocal tract is primarily influenced by the vertical position of the tongue, it is also modified by the displacement of the jaw which is, to some extent, independent of tongue position. Similarly, although the second formant is primarily determined by the horizontal position of the tongue, it is also influenced by the movements (especially rounding) of the lips (Stevens 1998). As a corollary to this, articulatory movements cannot be reconstructed solely on the basis of acoustic characteristics of the speech signal.

In Hungary, the creation of the MTA–ELTE Lendület Lingual Articulation Research Group (2016) arguably marks the beginning of a new era of research into articulation (and coarticulation). The establishment of new technical conditions opened up pathways of research that had formerly been impossible. The present study gives an overview of the background of research into articulation in Hungary and presents the tools and methods that can transform these investigations. We give an insight into the most important research carried out with these innovative technologies on the articulation of Hungarian adults.

## **2. Articulatory studies in Hungary – before 2000**

Articulatory investigations in Hungarian that are based on dynamic data instead of static images are scarce. The X-ray film technology (the so-called cineradiographic examination) was used by János Lotz in the 1960s (1966, 1967), Tamás Szende in the 1970s (1974) and Kálmán Bolla in the 1980s (1981b, 1981c) to study articulation characteristics of the Hungarian speech. Bolla analysed all the Hungarian vowels and consonants in his study. Five images of each speech sound were transmitted from the X-ray films to a computer and the computerized drawings were then phonetically analysed. In his studies, all the articulatory configurations were presented in drawings and the sizes of the vocal tracts were shown in tables. These data do not only help understand the mechanisms of Hungarian speech production and describe the basis of articulation but could also be used for a modern, articulation-based speech synthesizer. Bolla and his colleagues later gave a detailed description of the equipment used to make radiograms and the methodology applied to make the recordings (1986). According to this, the microcomputer technology was designed to embrace interlingual phonetic comparison. Bolla also experimented with the study of the lips (photolabiogram), the palate (palatogram) and the tongue (lin-

guogram). He did not only carry out articulatory research on Hungarian pronunciation (e.g. 1980, 1995), but also on other languages (Russian: 1981a; American English: 1981d; Finnish: 1985; German [with László Valaczkai]: 1986; Polish [with Éva Földi]: 1987).

Of the methods mentioned above, the X-ray technology is now considered to be obsolete. The most important reason for this is of ethical nature – in the modern scientific view, even voluntary exposure to harmful radiation for scientific purposes is regarded as improper. Furthermore, the use of X-ray machines requires special circumstances.

### 3. Articulatory studies with respect to Hungarian – the 2000's

Following the experiments in the 1980s, Hungarian articulation was not researched for a long time. As the necessary equipment was not available in Hungary, some sporadic investigations were conducted in foreign laboratories.

Beňuš and Gafos (2007) studied Hungarian vowel harmony with electromagnetic articulography and ultrasound tongue imaging, more specifically the phenomenon of transparent vowels (vowels that are neutral from the perspective of vowel harmony) /i i: ε e:/ not transmitting their quality to the vowel in the suffix. The authors explained the transparency of these vowels with their coarticulatory characteristics. Three participants were involved in the research and an ultrasound image was made of one of them. In the research material (target sentences), there were monosyllabic harmonic (e.g. *hír* ‘a piece of news’ – *hírek* ‘news’) and non-harmonic (e.g. *ír* ‘s/he writes’ – *írnak* ‘they write’) words and trisyllabic suffixed words, in which the final vowels of the stems were transparent (e.g. *bilivel* ‘with potty’ vs. *bulival* ‘with party’). The results showed a link between vowel harmony and articulatory characteristics as the tongue took a more retracted position during the pronunciation of non-harmonic stems than with harmonic stems. However, based on acoustic measurements, Blaho and Szeredi (2013) questioned the relevance of these results because they did not find any similar connection in the acoustic data. As a consequence, further research is needed in order to answer the question whether phonetic characteristics can account for the morpho-phonological behaviour of these vowels.

In 2008, Mády used electromagnetic articulography to examine Hungarian vowels in the normal and fast speech of two speakers. This research aimed at showing whether the articulatory characteristics (tongue displacement and jaw openness) are distinct in the case of phonologically short and long vowels pronounced in two different situations (normal and slow speech tempo). Mády reported a stronger coarticulatory effect during the pronunciation of short vowels, which is congruent with the results of research on German fortis (tense) and lenis (lax) vowels (Hoole–Nguyen 1999).

Recent investigations into Hungarian vowels were also conducted with electromagnetic articulography. Deme et al. (2016) examined the pronunciation of vowels sung at high fundamental frequency by a soprano singer. The focus of the investigation was on the mechanisms of the tongue and the jaw during the pronunciation of all the standard Hungarian vowels. The research was inspired by the results (primarily gained from acoustic and perceptual measurements) which showed that vowels in singing at a high fundamental frequency are distinct from their realizations in normal speech. The aim of the research was to record the articulatory mechanisms that underlie this phenomenon. According to the results of the EMA-analysis, the singer systematically changed the position of the tongue and the jaw (increased the angle of jaw openness and lowered the vertical tongue position) as the fundamental frequency of singing reached and then exceeded the  $F_1$ -value of vowels pronounced in normal speech. However, the lowering of the back of the tongue could already be observed below these critical frequencies. In addition, it

was found that below 988 Hz the singer achieved  $F_1 : f_0$  tuning by the unique combination of tongue and jaw movements specific to the intended vowel qualities, at 988 Hz the tuning was achieved by jaw opening and resulted in a uniform tongue and jaw position across all vowels.

In a subsequent study, Deme et al. (2017) analysed the vowels pronounced by three Hungarian and three German soprano singers at a high fundamental frequency. In their research, they compared all the standard Hungarian and German vowel qualities and used the same method for the recording of data as in the previous research. According to the data, the vertical tongue position lowered in each singer with raising of the  $f_0$ . The tongue position and the openness of the jaw systematically changed as the  $f_0$  reached and then exceeded the  $F_1$ -value of vowels in speech. The strategies of  $F_1 : f_0$  tuning were the following. (i) In the case of low  $f_0$ , the lowering of the vertical tongue position was observed during the production of close vowels. (ii) In the case of high  $f_0$ , especially at fundamental frequencies  $f''$  (698 Hz) and  $h''$  (988 Hz), the increase of jaw opening was observed during the production of more open vowels. (iii) This was, however, also accompanied by the lowering of the tongue. In addition, significant individual differences were recorded among the Hungarian and German participants regarding the retainment of articulatory differentiation of vowels. By contrasting the results of Hungarian and German singers, we concluded that there is no or only marginal dependence between the articulatory strategies of soprano opera singers to raise the  $f_0$  and the mother tongue of the singer, if the vowel systems of the languages we compare, have only minor differences.

#### 4. Recent articulatory studies with respect to Hungarian

The MTA–ELTE Lendület Lingual Articulation Research Group was founded in 2016. Its primary goal is to investigate coarticulation in Hungarian speech with articulatory devices. The following methods are available in our laboratory:

- (i) electromagnetic midsagittal articulometry (EMA), which is suitable for imaging a limited number of flesh-points;
- (ii) ultrasound tongue imaging (UTI), which gives information of the midsagittal view of the global tongue surface;
- (iii) electroglottography or laryngography (EGG), which is used for the measurement of the degree of contact between the vibrating vocal folds during voice production.

Without being exhaustive, research topics so far have been analysed are the following: the effect of prominence on coarticulation patterns (Markó et al. 2019b) and in relation to this, glottal marking on word and utterance-initial vowels (Markó et al. 2019a); articulatory timing of singleton, geminate and degeminated consonants and singleton consonants in clusters (Deme et al. 2019); voicing and tongue shapes in Hungarian singleton and geminate obstruents (Percival et al. 2020); phonation changes during emotion-inducing events (Bartók 2019); articulatory behaviour of the vowels of antiharmonic stems from the perspective of the horizontal position of the tongue (Markó et al. 2019c, d). In the present review we summarize the main findings of some of these studies with respect to articulation. Both a more detailed description of the above mentioned methods and exact measurement data (mean values, standard deviations, results of statistical analysis, and the presentation of the findings in figures) are presented in the papers referred to. Most of these papers are available online, and can be found on the research group's website: <http://lingart.elte.hu/en/publikaciok/>.

#### 4.1. Articulatory organization of geminates in Hungarian

Hungarian expresses semantic differences by using contrastive consonant phoneme length, see e.g. *ép* ‘healthy’ ~ *épp* ‘right now’. In theoretical works, duration is considered to be the main acoustic cue that makes the singleton-geminate phonological contrast in consonants. It is also traditionally assumed that geminates do not occur flanked by another consonant on either side, and that in these positions, geminates surface as short. This process is called degemination (Siptár–Törkenczy 2007).

On the basis of acoustic data, previous research concluded that in line with other languages that exhibit the contrast, it is indeed durational properties, especially closure duration, that are the most important correlates of the singleton-geminate opposition in Hungarian stops (Neuberger 2015; Olaszy 2006; Pycha 2009, 2010). Siptár and Grácz (2014) analysed some fricative and stop geminates in degemination cases, flanked by varying consonants. The authors concluded that among degeminated and singleton /t/ and /p/ realisations, singletons (in  $C_1C_2$ , either as  $C_1$  or  $C_2$ ) were the longest, followed by degeminated geminates (flanked by a  $C_2$  on one side), and singletons in  $C_1C_2C_3$  sequences (as  $C_2$  consonants).

In a study (Deme et al. 2019) we analysed several acoustic and articulatory features of singleton, geminate, and degeminated (voiceless) stops in Hungarian, to examine if (i) degemination neutralizes the singleton-geminate opposition in the acoustic and articulatory domain, (ii) singletons in  $C_1C_2$  clusters, and geminates in degeminating  $C_1C_1C_2$  positions differ in the extent of articulatory overlap they exhibit with a following heterorganic consonant, and (iii) slower tongue rise and longer preceding vowel duration is observable in geminates (compared to singletons), and if they are independent. For the articulatory analysis electromagnetic articulography was applied.

Consonant duration and total consonant cluster duration as measured in the acoustic signal, and the duration of the gestural plateau detected in the articulatory signal unanimously showed that degemination does not reduce stops to intervocalic singletons, but rather to singletons that are flanked by another stop consonant (i.e., singletons in two-term clusters). Articulatory data further suggests that degeminated stops and two-term clusters form an in-between category between geminates and singletons. As far as the timing of the articulatory gestures, more specifically, the articulatory overlap of gestural plateaus is concerned, we found that two-term clusters and degeminated stops differed only in lingual-labial (/pt/ ≠ /ppt/), but not in labial-lingual (/tp/ ≈ /tpp/) clusters, that is, degemination reduced geminates to singletons in C-clusters dependently of the place of articulation of the stops. Further, our results supported the findings of Fujimoto et al. (2015) showing that a preceding vowel does not show shortening (as one might expect) but lengthening before geminates. However, we also found the same trend for simple  $C_1C_2$  clusters. Moreover, we found a similarly slow tongue rise for both geminates and singletons in two-term clusters, which suggests that in some aspects, the phonetic implementation of geminate stops resembles that of two-term stop clusters. And finally, we found a strong correlation of tongue rise and preceding vowel duration, suggesting that preceding vowel duration may very well be considered a mere side effect of slower tongue movement in geminates and two-term clusters.

Even though Hungarian exhibits voiced geminates in a distinctive function, like in *megy* ‘go S3’ vs. *meggy* ‘sour cherry’ this pattern is not very frequent even in this language. Among various languages, moreover, voiced geminates are rather uncommon because of the articulatory difficulty of synchronously maintained voicing and obstruction. Voicing, therefore, has been found to vary in geminates in some languages (e.g. in Tokyo Japanese by Kawahara 2015). Although Hungarian acoustic research has found that voicing in singletons is variable

(e.g. Grácz 2013), this effect in Hungarian geminates has not been studied so far. In a study (Percival et al. 2020) we used EGG and UTI to investigate articulatory correlates to voicing in geminate as opposed to singleton consonants in Hungarian. With the help of EGG, we investigated whether voiced geminate obstruents are fully voiced in Hungarian, partially devoiced, or variable. Ultrasound can give an answer to the question if tongue position differ across singleton and geminate obstruents in Hungarian. In previous studies, coronal geminates were found to be produced with greater lingual-palatal contact and a higher and flatter tongue in various languages including Japanese, Korean, Italian and Oromo (Kochetov–Kang 2017; Payne 2006; Percival et al. 2019), suggesting that in geminates the tongue more fully reaches its targeted place of articulation than in singletons. These findings associate gemination with fortition.

We followed up on these studies by examining whether there is evidence for differences in lingual articulation in geminates compared to singletons in Hungarian. As previous studies concentrated on coronal stops, we additionally asked if similar patterns of tongue raising or fronting can be found for geminates at other places of articulation as this could indicate how closely the pattern is tied to gemination in general versus a tongue pull mechanism limited to coronals. Therefore, voiced and voiceless bilabial, alveolar and velar stops, and alveolar fricatives were involved both in singletons and geminates. We also examined the nature of the relationship between voicing and advanced tongue root, as previous research (e.g. Ahn 2018 for English and Brazilian Portuguese) has found advanced tongue root occurring with phonological but not necessarily phonetic voiced consonants.

We supposed that voiced obstruents are produced with advanced tongue root, an articulatory strategy which facilitates voicing. Given the articulatory difficulty in producing voiced geminates, we predicted partial or variable voicing in geminates and more use of advanced tongue root. However, EGG results did not show difference in voicing between singleton and geminate consonants. While voiceless consonants were generally voiceless, voiced consonants, both singletons and geminates, varied considerably in percent voicing. It seems that Hungarian geminates display variable behaviour similarly to what Grácz (2013) found for singletons in her acoustic study, they do not seem to be consistently semi-devoiced, or fully voiced.

As a function of voicing (based on phonological category: voiced and voiceless), ultrasound results did not differ, but some significant interactions with place of articulation and radius number in pharyngeal and velar regions were suggestive of advanced tongue root for many voiced obstruents. When phonetic voicing was included in addition to phonological voicing in the model, phonological voicing remained significant only in certain interactions, while percent voiced was a significant main effect. This is unexpected as it tentatively suggests that tongue root is better predicted by phonetic than phonological voicing in Hungarian, contrary to what Ahn (2018) found for devoiced stops in English. This may suggest that advanced tongue root is not automatically implemented as a strategy to enhance voicing in Hungarian. Follow-up research is needed to investigate the robustness of this finding with further analysis methods, and to compare it with other languages with variable and semi-voiced geminates.

#### **4.2. Articulatory analysis of transparent vowel /i:/ in antiharmonic Hungarian stems**

Backness harmony in Hungarian is a highly productive process, and due to the exceptional behavior of so-called neutral or transparent vowels, it has been analysed extensively in the phonological literature (see e.g. Hayes et al. 2009). Hungarian vowel harmony is stem-controlled, and operates in the left-to-right direction, i.e., the backness of the stem's final

vowel assigns the backness of the suffix vowel. Most of the suffixes show front-back alternation in Hungarian, and suffix vowels receive their [ $\pm$  back] quality from the [ $\pm$  back] quality of the adjacent stem-final vowel (Siptár–Törkenczy 2007).

In the phonological domain of the Hungarian vowel system harmonic and neutral vowels can be differentiated. Harmonic vowels can be classified as front, such as [y y: ø ø:], and back, as [u u: o o: ɒ a:]. In the case of alternating suffixes and harmonic stem final vowels, backness harmony governs the quality of the suffix, without exception, e.g. *ablak-ban* /ɒblɒkβɒn/ ‘window-loc’, *üst-ben* /yʃtβɛn/ ‘cauldron-loc’. Neutral vowels are phonetically front unrounded [i i: e: ɛ], but from the phonological aspect they are neither front nor back, as they are transparent with respect to harmony. If the stem final vowel is neutral/transparent, the backness of the suffix vowel is governed by the last harmonic vowel within the stem, e.g. *kastély-ban* /kɒʃte:jβɒn/ ‘castle-loc’.

The question thus arises whether a back or a front suffix is selected when the stem is monosyllabic, and its vowel is neutral/transparent. In Hungarian, both patterns can be observed. We can find stems selecting front suffixes (harmonic stems), where the phonetically front unrounded vowels [i i: e: ɛ] behave as phonologically front ones, e.g. *víz-ben* /vi:zβɛn/ ‘water-loc’, *kéz-ben* /ke:zβɛn/ ‘hand-loc’. However, other monosyllabic stems with these vowels are followed by back suffixes (antiharmonic stems), e.g. *sír-ban* /ʃi:rβɒn/ ‘tomb-loc’, *cél-ban* /tse:lβɒn/ ‘target-loc’.

In one of their experiments, Beňuš and Gafos (2007) analysed monosyllabic antiharmonic target words without any suffix but in carrier sentences. As the authors mentioned, they had tried to compile a set of stimuli in which the front- and back-selecting stems were comparable as much as possible, however, some of the surrounding consonants differed in their place of articulation, and these differences might have had an effect on the data. Therefore, in our study (Markó et al. 2019d), homophonous front-selecting (harmonic) and back-selecting (antiharmonic) stems (*nyír* /ɲi:r/ ‘birch’, ‘trim’ and *szív* /si:v/ ‘heart’, ‘suck’) were chosen, and electromagnetic articulographic experiments were conducted in order to test the hypothesis that in the back-selecting stems, the tongue was more retracted during the articulation of /i:/ than in the front-selecting stems. The target words were analysed both in isolation (isolation setup), where neither a suffix nor a carrier sentence were applied, and in carrier sentences (sentence setup), where the target word was positioned at the beginning of the sentence, and was followed by a word containing (i) only front vowels (*éppen* /e:p:ɛn/) or (ii) only back vowels (*ugyan* /ujɒn/). The horizontal position of four receiver coils (one on the tongue tip, one on the tongue blade, and two on the tongue dorsum) were obtained at the temporal midpoint of the target vowels. The results showed that neither the horizontal positions of the receivers nor the formant values varied as a function of the harmonicity of the stem in either the isolated or the coarticulated setup.

Based on these results, the conclusions formulated by Beňuš and Gafos (2007) on the sub-phonemic differences between the realizations of transparent vowels in front- and back-selecting stems are to be handled with care. On the basis of our data obtained with a well-controlled material, it seems reasonable to suggest that sub-phonemic differences (if they exist) cannot be traced back to (different) tongue positions associated with the transparent vowels’ realizations in front- (harmonic) and back-selecting (antiharmonic) stems.

### 4.3. Phonatory changes during emotion-inducing game events

According to appraisal models of emotion (e.g. Ortony et al. 1990; Scherer 2001), behavioral and physiological reactions to affective stimuli are a result of cognitive appraisal of the stimuli. Appraisal is described by the Component Process Model (CPM, Scherer 2001) as a process consisting of several subsequent Stimulus Evaluation Checks (SECs). The result of these steps of evaluation determines the emotional state and physiological reactions of the organism.

In a study (Bartók 2019), two such SECs, goal conduciveness and discrepancy from expectations were manipulated in a computer game, with the aim to describe their effect on vocal fold vibration. Speech was acquired during voice commands controlling the game, resulting in utterances that could capture the induced emotional effects right at the time they occurred. Hypotheses concerning these effects were formed based on the physiological changes predicted by the CPM for different results of these SECs and their possible effect on phonation, while also considering phonatory patterns observed in acted emotions, since such portrayals often build on representations of spontaneous emotional reactions. It was supposed that goal conducive game events that are congruent with the subjects' expectations lead to a more frequent occurrence of nonmodal phonation types, lower  $f_0$ , lower H1-H2 for females and higher H1-H2 for males relative to the subjects' emotionally neutral speech. However, goal obstructive game events and events discrepant from the subjects' expectations both were expected to lead to a less frequent occurrence of nonmodal phonation types, higher  $f_0$ , and higher H1-H2 relative to the subjects' emotionally neutral speech.

Phonatory variation was quantified in two ways: we determined phonation type manually, after which acoustic measurements were carried out on the modal parts of the analysed vowels. The two acoustic measures taken were fundamental frequency ( $f_0$ ) and the difference between the first two harmonics (H1-H2). H1-H2 is a measure well-suited to describe the degree of glottal constriction (Keating–Esposito 2007) as it correlates highly with the Open Quotient (OQ) (Shue et al. 2010), i.e., the proportion of a glottal cycle in which the glottis is open (Holmberg et al. 1995). Higher values of H1-H2 would suggest a more breathy phonation, while low values indicate irregular phonation.

Although no difference was shown in the frequency of phonation types between different results of the manipulated SECs, significant interaction effects of gender, discrepancy and conduciveness were found on the acoustic parameters measured on the modal parts of the voice commands. This could mean that emotions induced in this highly controlled, laboratory setting lead to subtle phonatory changes. The interaction effects of the two manipulated SECs and gender for both acoustic measures indicate that emotional reactions can only be captured in female speech. This could be explained by differences in the degree of emotional reactivity and emotion regulation between genders. Apart from the lay belief that females tend to be more emotional (Grossman–Wood 1993), several studies using physiological measures of emotional arousal and attention suggest that females are more reactive to emotional stimuli than males (e.g. Bradley et al. 2001; Grossman–Wood 1993; Kemp et al. 2004).

We found that for females,  $f_0$  is higher when facing game events that are discrepant from expectations, while congruent events lead to a decrease in  $f_0$ . This effect is likely to be caused by increased muscle tension when facing unexpected, discrepant stimuli and decreased tension in case of expected events (Johnstone et al. 2001). The effect is stronger in case of relaxation at goal conducive events.

We also found a lowering of H1-H2 values for discrepant, obstructive and congruent, conducive events in the phonation of females. H1-H2 lowering indicates a shift from females' habitually breathy phonation (Hanson–Chuang 1999) to a more modal one, as a result of the



predicted increase in overall muscle tension (Johnstone et al. 2001). In the case of conducive, congruent events, low H1-H2 together with the low  $f_0$  measured in this condition could mean that in this case, H1-H2 lowering does not simply indicate a more modal phonation, but rather a shift towards a more irregular phonation caused by relaxation, similarly to the frequent occurrence of irregular phonation when Hungarian females express contentment (Bartók 2018).

## 5. Possible applications based on articulatory data

Several applications might be proposed, and one of them has already started to be developed using speech data of Hungarian. Silent Speech Interfaces (SSI) are a revolutionary field of speech technologies, built on the main idea of recording soundless articulatory movements, and automatically generating speech from the movement information, while the original subject is not producing any sound (Denby et al. 2010). This research area has a large potential impact in a number of domains, including the development of communication aids for impaired people. Recently, novel methods have started to be developed for analysing and processing articulation (especially the tongue and the lips) during human speech production.

Our goals are to test and improve recognition-followed-by-synthesis and direct synthesis in the field of silent speech interfaces. For these, 2D ultrasound of the tongue and lip video are used to image the motion of the speaking organs. We use high-potential machine learning methods, including various deep neural network architectures. In order to achieve the above goals, we first recorded parallel speech and tongue-ultrasound data with multiple Hungarian speakers. Next, we performed articulatory analysis on that, modeled the articulatory-to-acoustic mapping in various ways, and are evaluating them in objective tests and subjective listening experiments. To fulfill the above goals, a multidisciplinary team was formed with expert senior researchers in speech synthesis, recognition, deep learning, and articulatory data acquisition (Csapó et al. 2018).

SSIs are still in an experimental phase, but several fields of use are predicted by the literature (see e.g. Denby et al 2010) from laryngectomized patients to providing privacy for cellular telephone conversations.

In speech therapy, articulatory devices can also be extensively used (see e.g. Cleland et al. 2015; Preston et al. 2017), as they are able to visualize fine motor behaviour which is unseen with a help of a mirror or video recording. The technique which is used in these applications is biofeedback, which means that the therapists use a kind of electronic tool to monitor and amplify body functions that may be too subtle for being available at a conscious level. Electronic instruments (like UTI or EMA) detect bioelectric signals and supply the subject via sensory modalities (auditory, visual, tactile, or a combination thereof). On this basis, the subject might be able to gain control over these specific body functions (Davis–Drichta 1980). Up until now, Hungarian speech therapy has only used this biofeedback method by relying on the acoustic domain of speech, with transformation of the acoustic signal to a visual output for patients with hearing impairment, e.g.

- Varázsdoboz: <http://lsa.tmit.bme.hu/products/speco.html>;
- Beszédmester: <http://www.inf.u-szeged.hu/projectdirs/beszedmester/>;
- Beszédszisztens: [http://www.jgypk.hu/mentorhalo/tananyag/az\\_ikt\\_alkalmazasa\\_a\\_gyogypedagogiaban\\_V2/1010\\_beszdszisztens.html](http://www.jgypk.hu/mentorhalo/tananyag/az_ikt_alkalmazasa_a_gyogypedagogiaban_V2/1010_beszdszisztens.html).

Now the articulatory methods, especially UTI, are also available for the therapy of motor sensory deficits.

### Acknowledgements

This paper was supported by the Thematic Excellence Program of ELTE Eötvös Loránd University, Budapest, Hungary, by the Bolyai János Research Scholarship of the Hungarian Academy of Sciences, and the ÚNKP-19-4 New National Excellence Program of the Ministry for Innovation and Technology.

### References

- Ahn, Suzy 2018. The role of tongue position in laryngeal contrasts: An ultrasound study of English and Brazilian Portuguese. *Journal of Phonetics* 71: 451–467.
- Bartók, Márton 2018. A gégeműködés variabilitása az érzelemkifejezés függvényében [Variability of laryngeal mechanisms as a function of emotion expression]. *Beszédkutatás* 26: 30–62.
- Bartók, Márton 2019. Phonatory changes during emotion-inducing game events: the effect of discrepancy from expectations and goal conduciveness. In: Calhoun, Sasha – Escudero, Paola – Tabain, Marija – Warren, Paul (eds.): *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019*. Canberra, Australia: Australasian Speech Science and Technology Association Inc. 57–61. <https://pdfs.semanticscholar.org/95cd/8bd4f400560dbd6b1425d05e098fdf1b94ae.pdf>
- Beňuš, Stefan – Gafos, Adamantios I. 2007. Articulatory characteristics of Hungarian ‘transparent’ vowels. *Journal of Phonetics* 35: 271–300.
- Bolla, Kálmán 1980. *Magyar hangalbum* [A phonetic conspectus of Hungarian]. Magyar Fonetikai Füzetek 6. Budapest: MTA Nyelvtudományi Intézet.
- Bolla, Kálmán 1981a. *A conspectus of Russian speech sounds. / Атлас звуков русской речи*. Slavistische Forschungen. Band 32. Budapest–Köln–Wien: Akadémiai Kiadó – Böhlau Verlag.
- Bolla, Kálmán 1981b. *A magyar hosszú mássalhangzók képzése (Kinoröntgenografikus vizsgálat számítógéppel)* [Articulation of Hungarian long consonants (Cineradiographic analysis with computer)]. Magyar Fonetikai Füzetek 7. Budapest: MTA Nyelvtudományi Intézet. 7–55.
- Bolla, Kálmán 1981c. *A magyar magánhangzók és rövid mássalhangzók képzési sajátosságainak dinamikus kinoröntgenográfiai elemzése* [Dynamic cineradiographic analysis of articulatory characteristics of Hungarian vowels and short consonants]. Magyar Fonetikai Füzetek 8. Budapest: MTA Nyelvtudományi Intézet. 5–62.
- Bolla, Kálmán 1981d. *Az amerikai angol beszédhangok atlasza* [A phonetic conspectus of American English]. Magyar Fonetikai Füzetek 9. Budapest: MTA Nyelvtudományi Intézet.
- Bolla, Kálmán 1985. *A finn beszédhangok atlasza* [A phonetic conspectus of Finnish]. Magyar Fonetikai Füzetek 14. Budapest: MTA Nyelvtudományi Intézet.
- Bolla, Kálmán 1995. *Magyar fonetikai atlasz. A szegmentális hangszerkezet elemei* [Atlas of Hungarian phonetics. Elements of segmental sound structure]. Budapest: Nemzeti Tankönyvkiadó.
- Bolla, Kálmán – Földi, Éva 1987. *A phonetic conspectus of Polish / Atlas dźwięków mowy języka polskiego*. Magyar Fonetikai Füzetek 18. Budapest: MTA Nyelvtudományi Intézet.
- Bolla, Kálmán – Földi, Éva – Kincses, Gyula 1986. *A toldalékcsoport artikulációs folyamatainak számítógépes vizsgálata* [Computer analysis of articulatory processes of vocal tract]. Magyar Fonetikai Füzetek 15. Budapest: MTA Nyelvtudományi Intézet. 155–165.
- Bolla, Kálmán – Valaczkai László 1986. *Német beszédhangok atlasza* [A phonetic conspectus of German]. Magyar Fonetikai Füzetek 16. Budapest: MTA Nyelvtudományi Intézet.
- Blaho, Szilvia – Szeredi, Dániel 2013. Hungarian neutral vowels: a microcomparison. *Nordlyd* 40(1): 20–40.
- Bradley, Margaret M. – Codispoti, Maurizio – Sabatinelli, Dean – Lang, Peter J. 2001. Emotion and motivation II: Sex differences in picture processing. *Emotion* 1(3): 300–319.
- Browman, Catherine P. – Goldstein, Louis M. 1986. Towards an articulatory phonology. *Phonology* 3: 219–252.

- Cleland, Joanne – Scobbie, James M. – Wrench, Alan A. 2015. Using ultrasound visual biofeedback to treat persistent primary speech sound disorders. *Clinical Linguistics & Phonetics* 29(8–10): 575–597.
- Csapó, Tamás Gábor – Gosztolya, Gábor – Grósz, Tamás – Tóth, László – Markó, Alexandra 2018. Némabeszéd-interfész nyelvultrahanggal (Beszédgenerálás artikulációs mozgás alapján) [Silent speech interface with ultrasound tongue imaging (Speech synthesis on the basis of articulatory movements)]. Presentation at *Beszédkutató* [Speech Research] conference, Budapest, 18–19. October 2018.
- Davis, Sylvia M. – Drichta, Carl E. 1980. Biofeedback: Theory and Application to Speech Pathology. In: Lass, Norman J. (ed.): *Speech and Language* 3: 283–308.
- Deme, Andrea – Bartók, Márton – Grácsi, Tekla Etelka – Csapó, Tamás Gábor – Markó, Alexandra 2019. Articulatory organization of geminates in Hungarian. In: Calhoun, Sasha – Escudero, Paola – Tabain, Marija – Warren, Paul (eds.): *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019*. Canberra, Australia: Australasian Speech Science and Technology Association Inc. 1739–1743.
- Deme, Andrea – Greisbach, Reinhold – Markó, Alexandra – Meier, Michelle – Bartók, Márton – Jankovics, Julianna – Weidl, Zsófia 2016. Tongue and jaw movements in high-pitched soprano singing: A case study. *Beszédkutató* 2016: 121–138.
- Deme, Andrea – Greisbach, Reinhold – Meier, Michelle – Bartók, Márton – Jankovics, Julianna – Weidl, Zsófia – Markó, Alexandra 2017. Tongue and jaw articulation of soprano singers at high pitch in Hungarian and German. Presentation at *International Seminar on Speech Production*. Tianjin, China, 16–19. October 2017.
- Denby, Bruce – Schultz, Tanja – Honda, Kiyoshi – Hueber, Thomas – Gilbert, James M. – Brumberg, Jonathan S. 2010. Silent speech interfaces. *Speech Communication* 52(4): 270–287.
- Fujimoto, Masako – Funatsu, Seiya – Hoole, Phil 2015. Articulation of single and geminate consonants and its relation to the duration of the preceding vowel in Japanese. In: The Scottish Consortium for ICPhS 2015 (ed.): *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: The University of Glasgow. Paper number 0070 retrieved from <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0070.pdf>
- Grácsi, Tekla Etelka 2013. *Zörejangok akusztikai fonetikai elemzése a zöngésségi oppozíció függvényében* [Acoustic phonetic approach to the voicing opposition of obstruents]. PhD thesis. Budapest: ELTE.
- Grossman, M. – Wood, W. 1993. Sex Differences in Intensity of Emotional Experience: A Social Role Interpretation. *Journal of Personality and Social Psychology* 65: 1010–1022.
- Hanson, H. M. – Chuang, E. S. 1999. Glottal characteristics of male speakers: acoustic correlates and comparison with female data. *The Journal of the Acoustical Society of America* 106(2): 1064–1077.
- Hayes, Bruce – Zuraw, Kie – Siptár, Péter – Londe, Zsuzsa 2009. Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85: 822–863.
- Holmberg, Eva B. – Hillman, Robert E. – Perkell, Joseph S. – Guiod, Peter C. – Goldman, Susan L. 1995. Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech and Hearing Research* 38(6): 1212–1223.
- Hoole, Philip – Nguyen, Noel 1999. Electromagnetic articulography. In: Hardcastle, William J. – Hewlett, Nigel (eds.): *Coarticulation. Theory, data and techniques*. Cambridge: Cambridge University Press. 260–269.
- Johnstone, Tom – van Reekum, Carien M. – Scherer, Klaus R. 2001. Vocal expression correlates of appraisal processes. In: Scherer, Klaus R. – Schorr, Angela – Johnstone, Tom (eds.): *Appraisal processes in emotion: Theory, methods, research*. Series in affective science. New York: Oxford University Press. 271–284.
- Kawahara, Shigeto 2015. The phonetics of sokuon, or geminate obstruents. In: Kubozono, Haruo (ed.): *The Mouton handbook of Japanese language and linguistics*. Berlin: Mouton Gruyter. 43–78.
- Keating, Patricia A. – Esposito, Christina 2007. Linguistic voice quality. *UCLA Working Papers in Phonetics* 105: 85–91.

- Kemp, Andrew H. – Silberstein, R. B. – Armstrong, Stuart M. – Nathan, Pradeep Jonathan 2004. Gender differences in the cortical electrophysiological processing of visual emotional stimuli. *NeuroImage* 21(2): 632–646.
- Kempelen, Farkas 1791/1989. *Az emberi beszéd mechanizmusa, valamint a szerző beszélőgépezének leírása* [The mechanism of human speech and the description of the author's speaking machine]. Budapest: Szépirodalmi Kiadó.
- Kochetov, Alexei – Kang, Yoonjung 2017. Supralaryngeal Implementation of Length and Laryngeal Contrasts in Japanese and Korean. *Canadian Journal of Linguistics* 62: 18–55.
- Lotz, János 1966. Egy magyar röntgen-hangosfilm és néhány fonológiai kérdés [A Hungarian X-ray film and some phonological issues]. *Magyar Nyelv* 62: 257–266.
- Lotz, János 1967. Hangos röntgenfilm-vetítés a magyar nyelv hangképzéséről [X-ray film on articulation in Hungarian]. In: Imre, Samu – Szathmári, István (eds.): *A magyar nyelv története és rendszere* [The history and system of Hungarian]. Nyelvtudományi Értekezések 58. Budapest: Akadémiai Kiadó. 255–258.
- Mády, Katalin 2008. Magyar magánhangzók vizsgálata elektromágneses artikulográffal normál és gyors beszédben [Electromagnetic articulography analysis of Hungarian vowels in normal and fast speech]. *Beszédkutatás* 2008: 52–66.
- Markó et al. 2019a = Markó, Alexandra – Grácsi, Tekla Etelka – Deme, Andrea – Bartók, Márton – Csapó, Tamás Gábor 2019. Megnyilatkozáskezdő magánhangzók glottális jelöltsége a szintaktikai pozíció és a magánhangzó-minőség függvényében [Glottal marking of utterance-initial vowels as a function of syntactic position and vowel quality]. *Beszédkutatás* 2019: 30–53. [http://ojs3.mtak.hu/index.php/beszku/article/view/1603/Marko\\_et\\_al\\_Beszku27\\_2019](http://ojs3.mtak.hu/index.php/beszku/article/view/1603/Marko_et_al_Beszku27_2019)
- Markó et al. 2019b = Markó, Alexandra – Bartók, Márton – Csapó, Tamás Gábor – Deme, Andrea – Grácsi, Tekla Etelka 2019. The effect of focal accent on vowels in Hungarian: articulatory and acoustic data. In: Calhoun, Sasha – Escudero, Paola – Tabain, Marija – Warren, Paul (eds.): *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019*. Canberra, Australia: Australasian Speech Science and Technology Association Inc. 2715–2719.
- Markó et al. 2019c = Markó, Alexandra – Bartók, Márton – Csapó, Tamás Gábor – Grácsi, Tekla Etelka – Deme, Andrea: Articulatory analysis of transparent vowel /i:/ in harmonic and antiharmonic Hungarian stems: Is there a difference? In: *Proceedings of Interspeech 2019*. Graz, Ausztria. 3327–3331. [https://www.isca-speech.org/archive/Interspeech\\_2019/pdfs/2352.pdf](https://www.isca-speech.org/archive/Interspeech_2019/pdfs/2352.pdf)
- Markó et al. 2019d = Markó, Alexandra – Bartók, Márton – Csapó, Tamás Gábor – Grácsi, Tekla Etelka – Deme, Andrea: Az /i:/ artikulációs és akusztikai sajátosságai harmonikusan és antiharmonikusan toldalékolódó tövekben [Articulatory and acoustic characteristics of /i:/ in Hungarian harmonic and antiharmonic stems]. *Nyelvtudományi Közlemények* 115: 233–254.
- Neuberger, Tilda 2015. Durational correlates of singleton-geminate contrast in Hungarian voiceless stops. In: The Scottish Consortium for ICPHS 2015 (ed.): *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: The University of Glasgow. Paper number 0422 retrieved from <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0422.pdf>
- Olaszy, Gábor 2006. *Hangidőtartamok és időszerkezeti elemek a magyar beszédben* [Sound durations and temporal patterns in Hungarian speech]. Budapest: Akadémiai Kiadó.
- Ortony, Andrew – Clore, Gerald L. – Collins, Allan 1990. *The cognitive structure of emotions*. Cambridge: Cambridge University Press.
- Payne, Elinor M. 2006. Non-durational indices in Italian geminate consonants. *Journal of the International Phonetic Association* 3: 83–95.
- Percival, Maida – Csapó, Tamás Gábor – Bartók, Márton – Deme, Andrea – Grácsi, Tekla Etelka – Markó, Alexandra 2020. Geminata as fortition? Articulatory data from Hungarian. Presentation at *LabPhon17 Conference*. Vancouver, July 6–8, 2020.
- Preston, Jonathan L. – Leece, Megan C. – McNamara, Kerry – Maas, Edwin 2017. *Ultrasound biofeedback sample videos and practice data* (Preston et al. 2017). ASHA journals. Fileset. [https://asha.figshare.com/articles/Ultrasound\\_biofeedback\\_sample\\_videos\\_and\\_practice\\_data\\_Preston\\_et\\_al\\_2017\\_/5150119](https://asha.figshare.com/articles/Ultrasound_biofeedback_sample_videos_and_practice_data_Preston_et_al_2017_/5150119)

- Pycha, Anne 2009. Lengthened affricates as a test case for the phonetics–phonology interface. *Journal of International Phonetic Association* 39: 1–31.
- Pycha, Anne 2010. A test case for the phonetics–phonology interface: gemination restrictions in Hungarian. *Phonology* 27: 119–152.
- Scherer, Klaus R. 2001. Appraisal considered as a process of multilevel sequential checking. In: *Appraisal processes in emotion: Theory, methods, research*. Series in affective science. New York: Oxford University Press. 92–120.
- Shue, Yen-Liang – Chen, Gang – Alwan, Abeer 2010. On the interdependencies between voice quality, glottal gaps, and voice-source related acoustic measures. In: *Proceedings of Interspeech 2010*. Makuhari, Chiba, Japan. 34–37. [http://www.seas.ucla.edu/spapl/paper/shue\\_interspeech\\_10.pdf](http://www.seas.ucla.edu/spapl/paper/shue_interspeech_10.pdf)
- Siptár, Péter – Grácz, Tekla Etelka 2014. Degemination in Hungarian: Phonology or phonetics? *Acta Linguistica Hungarica* 61: 443–471.
- Siptár, Péter – Törkenczy, Miklós 2007. *The phonology of Hungarian*. Oxford: Oxford University Press.
- Stevens, Kenneth N. 1989. On the quantal nature of speech. *Journal of Phonetics* 17: 3–45.
- Stevens, Kenneth N. 1998. *Acoustic phonetics*. Cambridge MA – London: The MIT Press.
- Szende, Tamás 1974. A magyar hangrendszer néhány összefüggése röntgenográfiai vizsgálatok tükrében [Some interrelations of Hungarian sound system with respect to cineradiographic analyses]. *Magyar Nyelv* 70: 68–77.