

Intervocalic voicing of Hungarian /h/

Andrea Deme^{1,2}, Márton Bartók^{1,2}, Tekla Etelka Grácz^{4,2},
Tamás Gábor Csapó^{3,2} & Alexandra Markó^{1,2}

¹Eötvös Loránd University / ²MTA-ELTE “Lendület” Lingual Articulation
Research Group / ³Budapest University of Technology and Economics /

⁴Research Institute for Linguistics, Hungarian Academy of Sciences

In this study, we investigated whether the amount of voicing, and the sound quality (expressed in HNR) of /h/ in syllable onset are affected by intervocalic context (vs. post-pausal context), by backness and openness of the flanking vowels, and by a phonological conditioner, pitch-accent. We showed that both measures were similarly high in all intervocalic positions (irrespective of the presence of word-boundary and pitch-accent), while after a pause they were substantially lower, meaning that /h/ was voiced, and more modally voiced intervocalically, than after a pause. Further, interaction of this effect with that of vowel features led us to conclude that open, close, back, and front vowel groups should be considered internally heterogeneous with respect to their effect on /h/ voicing.

Keywords: phonetic analysis, fricative voicing, glottal fricative, laryngeal fricative, fraction of voiced frames, HNR, voice leak, intervocalic voicing, breathy voice

1. Introduction

In the Hungarian consonant inventory, all but one of the obstruents occur in voiced and voiceless pairs. The only exceptional obstruent is the laryngeal fricative /h/, which has no voiced counterpart in the system, but may be phonetically realized as a non-contrastive voiced [ɦ] segment (Siptár 2001; Siptár & Törkenczy 2007).¹ In our research, we aim to examine this allophonic alternation of the laryngeal fricative from a phonetic point of view, in an attempt to shed more light on the phonetic and phonological factors that may facilitate or restrain the occurrence of [ɦ]

1. Note that while the phoneme corresponding to the segment(s) at hand is referred to as a voiceless velar fricative /x/ in current phonological descriptions (see Siptár & Törkenczy 2007), it is described as a laryngeal fricative /h/ in the phonetic literature (see Bolla 1995; Gósy 2004; Kassai 2005). Merely for the sake of simplicity and since the authors of the present paper come from the field of phonetics, the latter practice is adhered to in the present paper.

in Hungarian, and thus to test previous claims made both in the field of phonology and phonetics on this issue. As a first step in this process, the present study investigates the effect of intervocalic context (vs. post-pausal word-initial position), two vowel quality features, viz. vowel backness and openness, and a phonological conditioner, pitch-accent, on the ratio of voicing that occurs in /h/ in syllable-onset in laboratory speech. As a second aim, we also investigate the general sound (but not necessarily voice) quality of /h/ tokens with respect to their periodicity, and as a function of the openness and backness of the context vowels. For this purpose, we analyze an additional acoustic parameter, harmonics-to-noise ratio (HNR), which is traditionally assumed to reliably and informatively quantify the amount of voicing and turbulent noise in fricatives, and can account for subtle differences of modally voiced, breathy voiced, and voiceless realizations on a continuous scale.

1.1 Background

To Hungarian phonology, the asymmetric behavior of /h/ in voicing assimilation seems to be of the greatest concern, as /h/ triggers regressive devoicing, but does not undergo voicing before voiced obstruents (see e.g., Siptár 2001; Siptár & Törkenczy 2007). Its allophonic voiceless–voiced alternation, on the other hand, is assumed to be clear-cut: /h/ is taken to occur (almost) invariably as [h̥] in intervocalic positions (Siptár 2001; Siptár & Törkenczy 2007) (hereafter, in *VhV*) as well as in sonorant-*h*-vowel positions (hereafter, *sonhV*) (Siptár 1994/2016) due to what is often considered a purely phonetic process (Szigetvári 1998). Despite the prevalent view that emphasizes the phonetic nature of the intervocalic voicing of /h/, in some sources we also find suggestions that the occurrence of [h̥] is to some extent also conditioned phonologically. According to Siptár (1994/2016), in formal or slower speech styles /h/ may be realized as [h̥] only in intervocalic contexts where /h/ is followed by an unaccented vowel, as in *tehén* [tehe:n] ‘cow’. In these cases, additionally to voicing, /h/ may also be phonetically unrealized, resulting in such surface forms as [te:e:n]. If, however, /h/ is followed by an accented vowel, as in *a hír* [v'hi:r] ‘the news’, /h/ may not be voiced and cannot be deleted either (i.e., no *[v'i:r] may occur), or at least not in the speech style mentioned above. By contrast, as Siptár (1994/2016: 265) claims, in fast/casual speech, /h/ may be realized as voiced also in the *a hír* [v'hi:r] context, as in this speech style the rule governing the voicing of /h/ does not block voicing in the onset /h/ of an accented syllable.

To summarize the above phonological description of the behavior of Hungarian /h/ in syllable onset, we highlight three major claims: the realization of /h/ as a voiced [h] segment is facilitated by (i) intervocalic context, (ii) sonorant-vowel context, and (iii) the absence of prominence on the syllable starting with /h/. Such

claims could easily be supported or falsified by objective empirical data, and thus one might expect that we find such data in the literature. However, as our review of empirical phonetic research will show, this is not the case. The effect of the first two factors, i.e., intervocalic and sonorant-vowel contexts, has not yet been tested under strictly controlled experimental conditions, while the effect of the third one, namely the effect of prominence, has not yet been addressed in any kind of structured empirical studies at all.

Before we review the moderate amount of empirical evidence on the voicing of /h/ in Hungarian, we make a brief note here on how the very same phenomenon is treated in other languages. Voicing of /h/-like segments in intervocalic position, or in a voiced environment in general is claimed to be present in many languages: British English (Ladefoged & Johnson 2011: 157), Amharic (IPA Handbook 1999: 48), Czech (IPA Handbook 1999: 70), Hebrew (IPA Handbook 1999: 98), and Korean (IPA Handbook 1999: 112), to name a few. However, to the authors' knowledge, to date there are only two studies by Mitterer (2018) and by Teras (2018) which provided some quantitative data on the amount of voicing found in intervocalic /h/ or on the ratio of voiced variants among intervocalic tokens. However, these studies provide no data on /h/ tokens in other than intervocalic (e.g., post-pausal) positions, or on the acoustic characteristics of breathy voice in /h/ tokens, which could have served as reference for our present analysis.

While investigating acoustic correlates of the singleton–geminate opposition in Maltese, Mitterer (2018) incidentally found that intervocalically, more than 80% of the analyzed Maltese singleton /h/ tokens were “fully voiced” (i.e., voiced throughout the whole segment), while the remaining 20% showed “a relatively uniform distribution of voicing leaks from 10% to 80%” (2018: 35). From the data presented on figures we can also infer that while the average duration of singleton /h/ tokens was 150 ms, the average duration of voice leak was approximately 70 ms, resulting in a ratio of 47% of voicing in intervocalic /h/ on average. Teras (2018) studied pronunciation variants of the short intervocalic /h/ in Estonian, and labelled /h/ tokens manually in a binary manner (voiced vs. voiceless). She found that 70% of the intervocalically positioned /h/ tokens were realized as voiced, while an additional 6% percent labelled as unvoiced were also “partly voiced” (2018: 84). Although these results do not lend themselves to direct comparison with ours, since Mitterer presents only absolute voicing durations (besides the gross ratios of “voiced” tokens presented by both Mitterer and Teras), both studies are important for our present analysis, mainly for two reasons. First, these data provide evidence that intervocalic /h/-voicing occurs in several languages. Second, they also further establish the relevance of our present analysis by showing how limited the amount of available data is on the extent of voicing observable in the laryngeal fricative, not to mention the quality of voicing it exhibits.

In Hungarian phonetics, the status, place of articulation, and voicing characteristics of /h/ have long been a matter of debate. By now, the status of /h/ as a consonant has been clearly established in Hungarian phonetics (see Laziczius 1937, 1963/1979; Bolla 1995; Gósy 2004; Kassai 2005). As far as the allophones of /h/ are concerned, more specifically the place of articulation of these, and the contexts that facilitate the occurrence of each allophone, the phonetic handbooks of Hungarian show a diverse picture. Since the present study focuses on the voicing characteristics of /h/ in Hungarian, we do not detail the various allophones enumerated in these handbooks, rather we focus on the relevant ones: [h] and [ɦ]. All of the recent phonetic textbook authors (Bolla 1995; Gósy 2004; Kassai 2005) link voiced fricative [ɦ] to intervocalic contexts. Bolla (1995), on the basis of X-ray images taken from one male and one female speaker, considers these speech sounds pharyngeal, even though he uses the IPA symbols of voiceless and voiced laryngeal fricatives. With the exception of Bolla, [h] is undisputedly assumed to bear a glottal place of articulation by all of the cited sources. (Note, however, that due to strong coarticulation expected between /h/ and the neighboring vowels, the accompanying secondary oral gestures in /h/ warrant further discussion, see Section 1.3.) To summarize these views on the voicing feature of the above mentioned allophones, (i) /h/ is unanimously claimed to be voiceless in post-pausal syllable onset, and (ii) it is unanimously claimed to be voiced in syllable onset intervocalically. These observations lead us to the last but, with respect to the present study, most important point in our summary, the issue of voicing in Hungarian /h/.

In the first empirical studies which dealt with /h/ in Hungarian, with regard to its voicing, two important questions were formulated: (i) if the laryngeal fricative sounds may be voiced, and (ii) how one should interpret their uncommon voicing patterns, which combine two articulatory gestures: the incomplete adduction and the clearly observable undulation of the vocal folds. As Laziczius (1937) points out citing Meyer and Gombocz (1909), voiced [ɦ] was observed instrumentally already in the early 1900's between vowels (or between voiced sounds in general, as first formulated by Meyer, according to Laziczius) in 20 tokens using a kymograph (Meyer and Gombocz 1909, see Laziczius 1937). That is, there was evidence available to formulate the tentative conclusion that a voiced laryngeal fricative exists and occurs allophonically in intervocalic positions in Hungarian. Along with this finding, according to Laziczius (1937), the most intriguing questions the existence of this voiced allophone raised were (i) if [ɦ] should really be considered an allophone of the laryngeal voiceless fricative, as the constriction and frication produced in [ɦ] is fundamentally different from that observable in its voiceless counterpart [h], and (ii) if [ɦ] should be simply considered "voiced", as its voicing without complete vocal fold contact is clearly different from the voicing of vowels or other obstruents. Regarding the first question, Laziczius (1937) implicitly seems

to suggest that despite the apparent differences between the fricative gestures in [h] and [ɦ], these sounds are not to be distinguished more than as allophonic variants, as they do not reflect a linguistically relevant opposition in Hungarian.² As one might expect, today's mainstream phonetic and phonological description of Hungarian follows this tradition. As to the latter question, he suggests that instead of voicing, the term (*glottal*) *murmur* should be used which encompasses both the fricative gesture and the voicing gesture in the glottis, and distinguishes the voicing of [ɦ] from the full voice observable in voiced obstruents and vowels (Laziczius 1937). Nowadays, additionally to *glottal murmur*, the term *breathy voice* (see also e.g., Maddieson 1984; Stevens 1999) and *whispery voice* (Laver 1994) are also used for the unique voicing patterns /h/ sounds exhibit. However, not only the terminology but also the articulatory and acoustic parameters assigned to these terms may also vary with authors.³ (For an illustration of differences in the articulation of [h] and [ɦ], see Esling 2005.)

In sum, studies in Hungarian phonetics from the first half of the 20th century already claimed that there is evidence to assume the existence of a voiced laryngeal fricative in Hungarian, and they also realized the peculiarity of the voice quality it displays. However, the systematic exposure of the factors that facilitates the occurrence of [ɦ], and the acoustic description of its voicing characteristics were at that time, and as we will see below, are still, missing from the literature.

As we already briefly discussed, in the phonetic textbooks published in the last decades, a very strong agreement is reached with respect to the occurrence of breathy voiced /h/, irrespective of its place of articulation. According to this consensus, and in line with phonology, Hungarian /h/ is (breathy) voiced in all *VhV* contexts (Laziczius 1963/1979; Bolla 1995; Kassai 2005), as well as in all *SonhV* contexts (Laziczius 1963/1979; Kassai 2005). But interestingly, these prevalent assumptions are treated as trivial, despite the fact that ever since the study of Meyer and Gombocz (1909) who observed the breathy voiced [ɦ] in only 20 intervocalic tokens of /h/ in unsystematically varying contexts, there has been only one experimental attempt to gather some more evidence for it, by the use of modern, and (at least in some sense) more reliable techniques, in more speakers, and via the analysis of more than 20 data points (Gósy 2005, see below). Moreover, the

2. Actually, the phonemic distinction between [h] and [ɦ] is generally very rare, as it is attested only in two of the 317 languages included in the UCLA Phonological Segment Inventory Database (Maddieson 1984).

3. *Breathy voice* and *whispery voice* are compound phonatory modes involving two particular types of voiceless airflow, *whisper* and *breath*, and a voicing gesture (for further discussion see Laver 1994; Gobl & Ní Chasaide 1995).

hypothesis that *SonhV* contexts also facilitate the voicing of /h/ to the same extent as *VhV* contexts, has never been addressed and corroborated experimentally. The same also applies to Siptár's (1994/2016) suggestion that prominence may play a role as a governing factor in the [ɦ] ~ [h] alternation.

As the first, and to date the only follower of the path Meyer and Gombocz had set for research, Gósy (2005) analyzed 50 intervocalic realizations of /h/, by the use of visual inspection of the spectrogram and the oscillogram, and a categorical analysis that differentiated only between *voiced* and *voiceless* realizations (the criterion for /h/ to be classified as voiced was whether it showed low-frequency periodicity for minimally two thirds of its length). The material consisted of a number of different Hungarian vowel qualities serving as *VhV* contexts in Hungarian words. As the aim of the study was an initial exploration of the occurrence of the voiced allophone of /h/, these *VhV* contexts varied unsystematically: some contexts were symmetrical, some asymmetrical, some occurred word-medially, while others at word boundaries. The vowel opening and backness features were also not systematically controlled factors, or at least their variation did not result in a numerically well-balanced design, and neither were these features tested in otherwise completely well-matched comparisons (e.g., vowels were not matched for the backness feature, when the effect of openness was evaluated). In the final analysis, Gósy (2005) concluded that /h/ is likely to undergo voicing (i) if the speech rate is fast,⁴ and the duration of the consonant is reduced accordingly, and (ii) if the /h/ sound is produced in the symmetric context of front vowels (as opposed to back vowels). More specifically, among all front contexts approx. 75% of /h/ tokens were realized with 100% voicing, while among back contexts, only approx. 38–39% (Gósy 2005: Figure 13).

Even though it may be claimed that due to the unbalanced speech material, the cited conclusions should be regarded as only tentative, they point to a very important question, namely, if the features of the vowels that serve as context for /h/ may play a role in facilitating the intervocalic voicing of /h/. More specifically, on the basis of the results of Gósy (2005), we may propose that both backness and openness should be tested as potential factors affecting the voicing of intervocalic /h/. The reason to include also openness in the analysis is basically the same as the rationale behind measuring the effect of backness, i.e., the interaction between the tongue position and the laryngeal settings that is exerted via a complex system of physiological and aerodynamical linkages between the tongue and the larynx structures.

4. Gósy (2005) provides no specific measure of 'fast' speech.

As articulatory studies tend to suggest, front close unrounded vowels (e.g., /i/ /e/) (and according to some sources, even the open front /a/, see Demolin et al. 2002) have a higher larynx position than back close, close-mid or open rounded vowels (as e.g., /u/ /o/) (Painter 1976; Hess 1998; Hoole & Kroos 1998; Demolin et al. 2002).⁵ Although the results are somewhat inconclusive for the effect of the backness feature alone (while Painter 1976 found implications of the backness effect via electromyography, Hoole and Kroos (1998) showed no clear support for it but found high interspeaker variability via digital video filming of the thyroid prominence), and are not really explicit about the effect of vowel openness alone (see e.g., Demolin et al. 2002), we may tentatively claim that front vowels display a higher position of the larynx than back vowels, while to a smaller extent, close vowels are also differentiated by higher positions of the larynx from low vowels. Moreover, larynx position also affects the production of voicing, since lower positions of the larynx induce a greater abduction of the vocal folds (Zenker & Zenker 1960; Pabst & Sundberg 1992), while higher positions increase adduction, and thus tension in the vocal folds (Sundberg & Askenfeld 1981; Honda et al. 1999). As a result, a higher position of the larynx also decreases the possibility of creating a glottal chink needed for the production of breathy voice or voicelessness.

Now, as a result of such governing forces of articulation as ‘articulatory economy’, i.e., the (maximal) reduction of metabolic costs of speech production (see e.g., Lindblom 1990 for further details), we may assume that the vertical position of the larynx may be lower in /h/ if it is flanked by back/open vowels symmetrically (rather than by front/close vowels), since back/open vowels also induce lower position of the larynx. This interrelation may be regarded as ‘secondary’ coarticulation, as coarticulation in this case affects articulatory gestures which are not traditionally considered to be primary gestures in the production of speech sounds, such as vertical larynx position. Consequently, we may also expect that the maintenance and/or the quality of voicing in /h/ varies as a function of vowel quality, especially vowel backness and openness. On this basis, we argue that the vocalic features openness and backness, affecting the vertical position of the larynx both in the vowel and possibly in the enclosed consonant due to coarticulation, should both be tested as potential factors that may also affect the control and the quality of voicing in the enclosed laryngeal fricative.

5. Actually, this interrelation is only true for languages that do not use the advanced tongue root [ATR] feature, since in the [ATR] languages the opposite relationship was observed empirically between vowel openness and larynx height, as reviewed and pointed out by Hess (1998).

1.2 Aims

Before discussing the aims of the present study, let us briefly recapitulate the claims made in the literature whose validity is at issue. First, it is claimed without firm empirical evidence that Hungarian /h/ in syllable onset is (breathy) voiced in all *VhV* contexts (Laziczius 1963/1979; Bolla 1995; Siptár 2001; Siptár & Törkenczy 2007; Kassai 2005), as well as in all *SonhV* contexts (Laziczius 1963/1979; Siptár 1994/2016; Kassai 2005). The most reliable experimental attempt addressing the effect of *VhV* context so far concluded that /h/ is likely to undergo voicing if it is produced in the symmetric context of front vowels (rather than back vowels) (Gósy 2005). However, the evidence supporting this conclusion was obtained in a study where the analyzed contexts were most probably not well-controlled for all features of interest, and thus the conclusions need further support and clarification. Moreover, on the basis of the findings of Gósy (2005), and the complex interlinks between lingual and glottal articulatory structures, we also propose that the effect of vowel openness should also be tested as a potential conditioner of /h/ voicing. Second, it is also claimed without objectively gathered empirical data that /h/ may be realized as [h̥] only in intervocalic contexts where /h/ is followed by an unaccented vowel (Siptár 1994/2016). For this claim we find no objective evidence in the literature.

In the present study, we aim at gathering empirical evidence for all but one of these claims and proposals. Our primary goal is to test four effects that may potentially affect the voicing characteristics of intervocalic /h/: (i) intervocalic context (vs. post-pausal word-initial position), two vowel quality features of the flanking vowels, (ii) backness, and (iii) openness, and (iv) a phonological conditioner, prominence.⁶ In the present study we test these effects in syllable-onset (i.e., the position of /h/ in the syllable is a controlled factor in the analysis), since in this position, the place of articulation of /h/ is undisputedly laryngeal (with accompanying secondary oral gestures due to lingual coarticulation; for more details, see Section 1.3).

On the basis of the findings of Meyer and Gombocz (1909), and Gósy (2005), and the previous professional experience of the authors with the segment at hand,

6. Note that due to reasons of space, this study is only concerned with intervocalic context, and symmetrical vowel contexts specifically. The latter decision is further motivated by the fact that the predictions are much clearer for symmetric contexts than for asymmetric contexts (see Gósy 2005), where further movements of the larynx are also expected during the production of /h/, i.e., a dynamic instead of a static production is expected, with less clearly predictable effects on phonation. However, we already started to gather and analyze data on the effect of the *sonhV* context, on which preliminary results were already presented elsewhere (Deme et al. 2018).

we hypothesized that intervocalic contexts invariantly elicit the voicing of /h/ in syllable onset, irrespective of the presence or absence of a word boundary and a corresponding sentence level prominence. However, in accordance with the findings of Gósy (2005), the phonetic characteristics of breathy voice, and the above outlined interaction of articulatory structures, we also expected that front and close vowels facilitate the maintenance of voicing in /h/ more (i.e., throughout the whole segment).

As a second aim, we also intended to give an initial, exploratory description of the quality of breathy voice Hungarian /h/ is claimed to exhibit, or to put it more precisely, to describe the quality of the sound (but not necessarily just the quality of the voice or voiced fraction of the sound) produced during the production of /h/ in intervocalic positions as opposed to the sound quality (may it be voiced, unvoiced, or partially voiced) it exhibits in a baseline post-pausal position. To grasp this general “sound quality” or signal characteristic difference between realizations, we measured and analyzed harmonics to noise ratio (HNR) as a function of the quality of the flanking vowels, with special attention paid to backness and openness. This acoustic parameter is briefly introduced in the next section, along with the acoustics of breathy voice known from the literature. As far as the quality of the produced sound, i.e., signal noisiness/periodicity of /h/ is concerned, we predict that higher and more fronted positions of the tongue (and the larynx) facilitate the increased adduction of the vocal folds, and thus reduce the breathiness, i.e., the amount of noise in /h/.

We find it important to emphasize here that to reliably test our hypotheses, the most fruitful approach is to use a phonetically well-controlled and numerically also well-balanced speech material, and a fine-grained phonetic analysis, which could account for voicing (and the quality of the signal in /h/ tokens in general) in a gradual way by quantifying these features on a continuous scale. In order to fulfill these expectations, in our experiment we used read laboratory speech, and measured the ratio of the voiced part (RVP) to express the amount of voicing, and harmonics-to-noise ratio (HNR) to express the quality of voicing in the segment at hand. Our approach in all of these respects constitutes novelty both in the topics of fricative voicing (especially in the laryngeal fricative) and the quality of breathy voice in general.

1.3 The acoustics of breathy voice, and acoustic parameters that quantify voice quality in fricatives

As discussed in the previous sections, the intervocalic allophone of Hungarian /h/ in syllable onset is claimed to be a voiced laryngeal fricative [h̥]; however, as we have also seen in detailed phonetic argumentations, the voicing of /h/ is

a very special type of phonation, namely *breathy voice*. Therefore, to reveal if intervocalic /h/ is really “voiced”, we must somehow grasp the (subtle) acoustic difference between the *voiceless* (or, applying Laver 1994’s terms, the *breath phonated*) [h] and the *breathy voiced* [ɦ] variant of the laryngeal fricative. Based on the acoustic and articulatory features associated with breathy voice (and summarized briefly below), we opted for the harmonics to noise ratio (HNR) parameter to operationalize this difference, which expresses the harmonicity/noisiness of the voice source.

In the production of breathy voice, the vocal folds vibrate very inefficiently, and they never come fully together during the “closed” phase. As a result, a substantial amount of air is constantly escaping through the glottis during the whole phonatory cycle in the form of a turbulent air stream, causing a clearly audible frication noise characterizing this special voice quality (see e.g., Laver 1994; Gobl & Ní Chasaide 1995; Stevens 1999). In accordance with this, and according to Gobl and Ní Chasaide (1995), an important acoustic cue that differentiates *breathy voice* from other “noisier” types of phonation, including voicelessness, is a dominant periodic component relative to the noise component. Consequently, we propose that HNR, a measure of signal harmonicity is appropriate to quantify the fine-grained phonetic difference between breathy voiced and voiceless realizations of /h/, and it may also reliably reflect the dominance of noise/periodic components within the groups of phonated variants. As its name suggests, HNR expresses the relation between periodicity and noise in a signal, and so it can be a measure of both degree of voicing and “noisiness”, i.e., how periodic (as opposed to aperiodic) a sound is (Gradoville 2011; G. Kiss 2013). A HNR of 20 dB means that 99% of the energy of the signal is in the periodic part, and 1% is noise; in other words, HNR approximating to 20 dB indicates a modally voiced (e.g., a vowel) segment, while a HNR of 0 dB means that there is equal energy in the harmonics and in the noise (see the Praat manual) (for details of calculation, see Boersma 1993). We argue that for the purposes of the present study, HNR is the most suitable parameter, since unlike other source parameters (e.g., spectral tilt or formant bandwidth based parameters, see Garellek to appear), it can be measured both in (partially) voiced and voiceless realizations, and can thus quantify the differences between these qualities on one single dimension. We must also emphasize here again, however, that there are no formerly established HNR values for the possible realization of the laryngeal fricative which could be used as a reference in the present analysis.

As a second important feature of the laryngeal fricative, we must also briefly mention here its susceptibility to coarticulation, which is also a reason why e.g., measures of spectral balance (e.g., center of gravity, COG) were not used here as a measure of voicing, despite its widely accepted use for this purpose in

fricatives (see e.g., Gradoville 2011; G. Kiss 2013). Since the laryngeal fricative is specified only for a constriction in the glottis and is underspecified for oral configuration (see Keating 1988: 282–283; Beckman 1995: 212), it can essentially be produced with any kind of anticipating or perseverating supraglottal (co)articulatory maneuvers that the economical production of the following or the preceding vowels may demand. Therefore, even in cases where the main place of articulation of /h/ is in the larynx (as is expected in the case of /h/ in syllable onset in Hungarian), there is most probably a secondary constriction also in the oral cavity when /h/ is preceded, followed, or enclosed by vowels. And this constriction inevitably shifts the center of gravity of the spectrum to a certain extent.

2. Methods

2.1 Participants

We recorded 19 speakers (11 male, 8 female) aged between 22 and 38 (average: 30). The speakers were monolingual native speakers of Hungarian with no speech or hearing disorders reported. Informed consents were collected from each participant before the recordings.

2.2 Material

We analyzed /h/ realizations in syllable onsets of Hungarian words embedded in meaningful sentences which we recorded as part of a bigger set of linguistic material. Tokens of /h/ were recorded in three conditions which varied in vowels as contexts, in the position of /h/ in the word, and in the syllable having or not having pitch-accent.⁷ The three conditions were the following. (i) Word-medial /h/ as the onset of an unaccented syllable (referred to as *VhV* in the analysis), e.g., *Teher* /teher/ *nyomta a vállát, nem a semittevés könnyűsége* ‘it was burden that she carried on her shoulders, not the unbearable lightness of doing nothing’. (ii) Word-initial /h/ in a pitch-accented word (the pitch-accent was elicited by pre-verbal focus position) as the onset of the accented syllable (referred to as *V#hV* in the analysis), e.g., *Keze hevesen* /keze#heveʃɛn/ *kutatott a nő után a sötétben* ‘his hands searched

7. In Siptár (1994/2016) it is not specified if the effect of ‘prominence’ should be regarded as an effect of lexical stress or sentential prominence, i.e., pitch-accent. Therefore, in the present study we opted for testing the stronger prosodic phrase boundary marker, pitch-accent, which thus should also exert the strongest effect (if prominence does play a role in the issue at hand).

for her vehemently in the dark' (iii) Word-initial utterance-initial (post-pausal) /h/ as the onset of the pitch-accented syllable (referred to as #hV in the analysis), e.g., *Herendi* /hɛrɛndi/ 'from Herend (name of a city)'. (Note that items in the #hV condition were one-word sentences, and while reading these sentences, participants were repeatedly reminded to read each item as an individual sentence.) Target items were collected by means of the 'szószablya' webcorpus (Halácsy et al. 2003). In this setting, condition (iii) served as a baseline, since in this condition /h/ is predicted to exhibit no voicing both by phonetic and phonological studies. As mentioned above, to reduce the number of variables, but still provide a representative and well-balanced set of vowels for the investigation of the effect of backness and openness, in the present study, we limited the analysis to symmetric V-contexts, where we used /u v i ε/ in all conditions systematically. In other words, we compared /h/ realizations in /uhu/, /vɥv/, /ihi/ and /ɛhɛ/ sequences, to test the effect of vowel backness (back /uhu/ and /vɥv/ vs. front /ihi/ and /ɛhɛ/) and vowel openness (close /uhu/ and /ihi/ vs. open /vɥv/ and /ɛhɛ/). Note that the feature of lip rounding covaried with backness (front vowels were both unrounded, while back vowels were both rounded), and showed a 50–50% distribution both in the open and the close groups (one rounded and one unrounded in each group). Since there is no open rounded front vowel in the Hungarian vowel inventory, this solution was the only one that could efficiently control for the effect of lip rounding (or at least balance it among the conditions). However, we are also aware that it potentially also exaggerated the effect that vowel backness may have exhibited, as lip rounding further lowers the position of the larynx (see Hoole & Kroos 1998).

To reduce the overall duration of recording sessions, we recorded 3 repetitions from each stimulus. Recordings were made in a sound-attenuated room with an omnidirectional microphone and an external sound card. The items were presented on a computer screen in a randomized order; the recording sessions were administered by the SpeechRecorder software (Draxler & Jänsch, 2004). After the exclusion of mispronounced tokens, a total of 1446 stimuli were analyzed.

2.3 Measurements

2.3.1 *Estimation of the voiced part*

As already mentioned, to quantify the extent to which /h/ was realized as voiced on a ratio scale (instead of using the categorical 'voiced' vs. 'voiceless' labels), we used the ratio of the voiced part to the total duration measure (RVP). For this purpose, we measured the total segment duration and the duration of voicing (more specifically, the fraction of locally unvoiced frames) in it. To find the optimal way for the detection of voicing, and thus the estimation of RVP in the target segments, we ran a pre-test on three speakers' data, i.e., on 450 stimuli. First, two authors of the

study segmented and labelled the target VCV and target CV segments, and they also annotated the voiced fraction of the /h/ realizations manually by the use of acoustic cues in the low-pass filtered acoustic signal (upper limit of the filter: 500 Hz). After the extraction of RVP in both annotations, we compared the data from the two annotators, and found close agreement between them ($\rho = 0.9, p < 0.001$). Then, we also extracted the RVP by the use of the voice report function in Praat (Boersma & Weenink 2017) using the settings suggested by Eager (2015): we used gender-specific f_0 -ranges of 70–250 Hz for males and 100–300 Hz for females. Finally, we compared the data of the second annotator and the RVP extracted by Praat. The comparison showed a very close agreement between the data extracted manually and automatically ($\rho = 0.7, p < 0.001$). On this basis, and based upon other independent sources which concluded that the voice report function in Praat may be regarded as statistically equivalent to manual annotation (see Eager 2015), we finally opted for manual annotation of the target segments but the use of automatic voice detection (i.e., Praat’s voice report) for the estimation of RVP in all participants. Typical examples of /h/ tokens in #hV, V#hV, and VhV conditions, and one atypical (i.e., almost voiceless) example of /h/ in VhV condition are illustrated in Figure 1 by the recorded audio waveform, spectrogram, and their corresponding manual segmentation. RVP data were also analyzed in interrelation with total token duration.

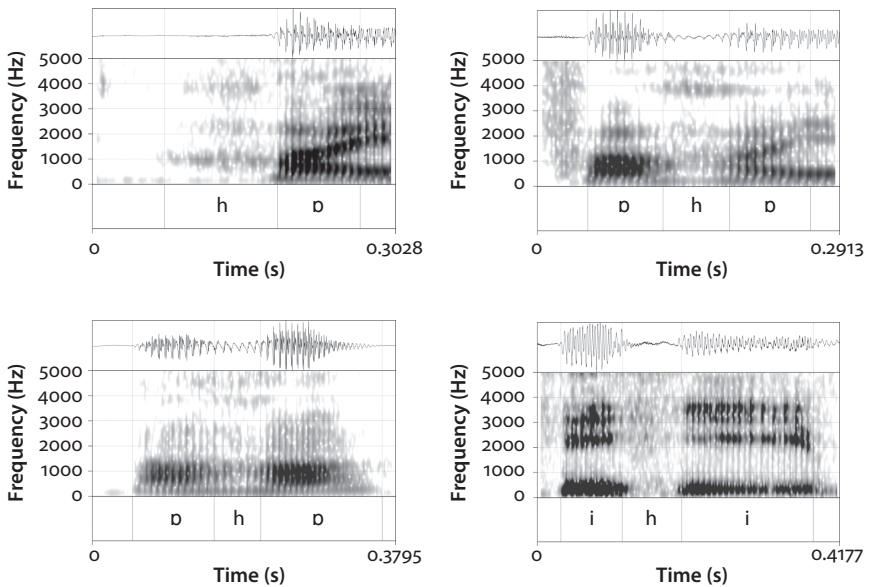


Figure 1. Upper row: typical examples of /h/ tokens in the #hV (left, voiced in 18%), and V#hV (right, voiced in 100%) conditions; bottom row: typical (left, voiced in 100%) and atypical (right, voiced in 25%) examples of /h/ tokens in the VhV condition

2.3.2 Acoustic measure of signal characteristics

Our second aim was to also investigate how intervocalic (vs. post-pausal) position affects the sound quality, or more specifically, the periodicity/noisiness of /h/. Moreover, we also intended to test if this feature of the /h/ realizations is also affected by the backness and openness of the neighboring vowels. For this purpose, we measured HNR, an acoustic parameter (which we also briefly introduced in Section 1.3) in Praat: we measured the mean HNR within the whole /h/ segment (HNR averaged over all frames, may it be voiced or voiceless), with the minimum pitch set to 100 Hz for women and 70 Hz for men (on the basis of Eager 2015), and the values of silence threshold, time step, and periods per second set to Praat's standard values. Note that HNR averaged for the whole segment is affected both by the amount of voicing present in the segment, and the intensity of the noise components. Therefore, this method has the advantage of encompassing both features of interest in one simple dimension, while it also has the disadvantage that it puts limits to the interpretation of the parameter at hand, as it does not simply reflect the quality of voicing in the voiced fraction of the sound, but rather an overall quality of the whole segment.

2.3.3 Statistical analyses

We used the lme4 package (Bates et al. 2015) in R (R Core Team 2017) to perform a linear mixed effects analysis of the effect of *vowel backness*, *vowel openness* and *condition* (as fixed effects) on RVP and HNR values. In these models, we also included random intercepts for subjects, and random slopes by subject for all fixed effects. Assumptions of homoscedasticity and normality were tested by visual inspection of residual plots. *P*-values were obtained via the Satterthwaite approximation available in the lmerTest package (Kuznetsova et al. 2017). Pair-wise comparisons were carried out with Tukey's post hoc tests available in the lsmeans package (Russell 2016). For graphical representation, we summarized the data with adjustment of the confidence intervals by removing inter-subject variability using the method proposed by Morey (2008). Correlation analyses were performed by Spearman's test.

3. Results

In general, 32% of /h/ tokens were realized with 100% RVP in all conditions, among which 97% were found in the two types of intervocalic conditions. Among tokens of intervocalic positions, 81% were realized as fully voiced. Among all /h/ tokens between front vowels 87% were realized as fully voiced (i.e., RVP = 100%), while among all tokens between back vowels, 75%. Overall RVP data (Figure 2, upper

left) shows that all the intervocalic conditions pattern together and are clearly distinguished from the baseline condition with a relatively high ratio of voicing. The group means and their standard deviations were as follows: #hV: $35 \pm 21\%$, V#hV: $88 \pm 23\%$, VhV: $96 \pm 13\%$. Figure 2 also shows the data broken down by vowel quality (upper right panel). Generally speaking, the tendencies in terms of all vowel qualities are highly similar to what is observable in the pooled data, that is, on this basis no notable vowel quality effect may be suspected. However, we also see that in the word-initial pitch-accented (V#hV) condition, the extent of voicing in the back /v/ ($78 \pm 27\%$) and /u/ contexts ($87 \pm 24\%$) is slightly lower than in the front vowel contexts ($94 \pm 18\%$ for both / ϵ / and /i/), while in the post-pausal syllable onset (#hV), the front /i/ ($29 \pm 19\%$) and / ϵ / ($34 \pm 20\%$) elicited a lower ratio of voicing in /h/ as opposed to the back /u/ ($37 \pm 22\%$) and /v/ ($39 \pm 20\%$).

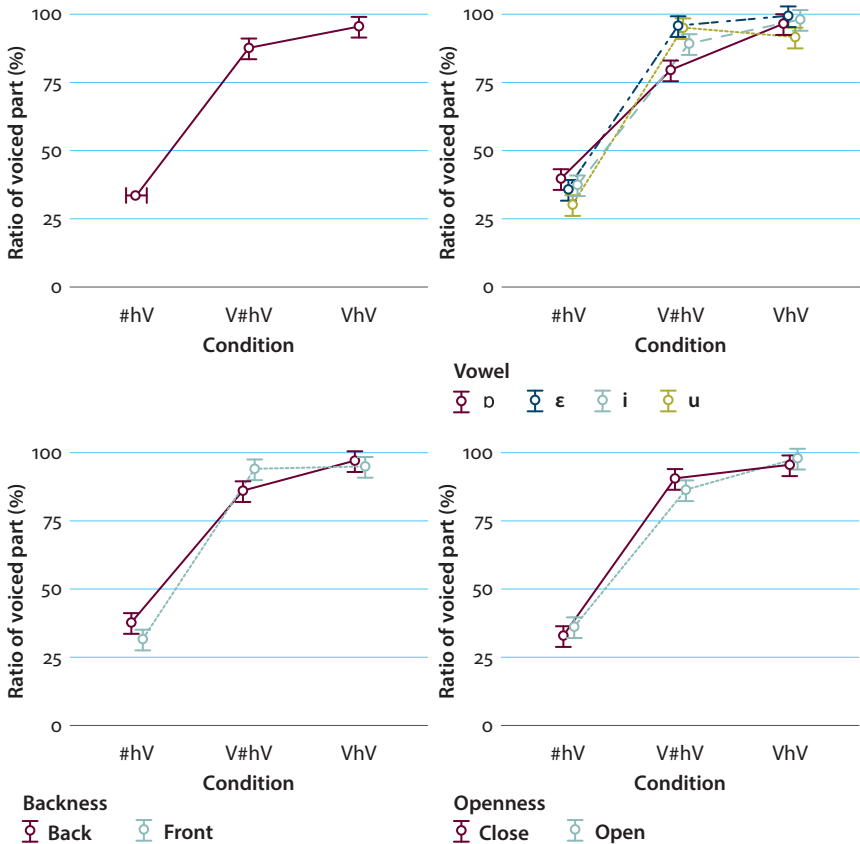


Figure 2. RVP in all /h/ segments pooled (upper left), and as a function of the right vowel's quality (upper right; /v/: black solid line; / ϵ /: black dotted-dashed line, /i/: grey dashed line, /u/: gray dotted line), vowel backness (bottom left), and vowel openness (bottom right) (mean \pm 95% confidence interval, adjusted for inter-speaker variability)

The above observations were partially corroborated by statistical analysis: according to a linear mixed model including the fixed factors *openness*, *backness*, and *condition*, and *speaker* as random effect, there was a significant interaction effect between *backness* and *condition* ($F(2, 1042.50) = 18.13, p < 0.001$) (Figure 2, bottom left), between *openness* and *condition* ($F(2, 1041.47) = 3.92, p = 0.02$), and between *openness* and *backness* ($F(1, 1113.64) = 8.79, p = 0.03$) on RVP (Figure 3).

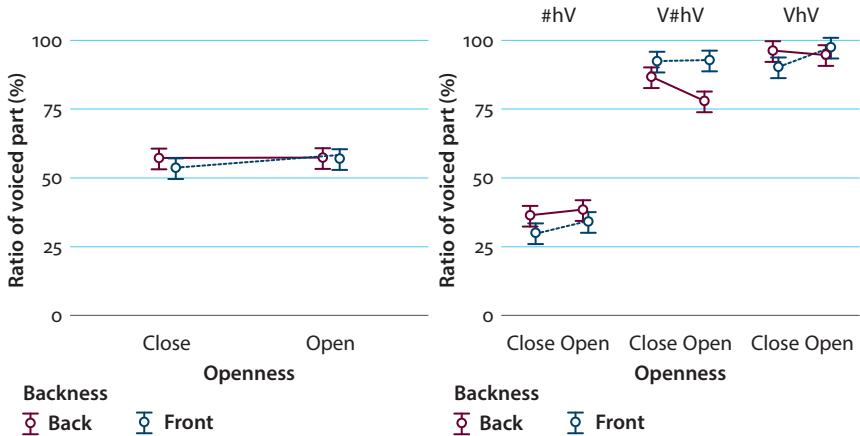


Figure 3. RVP as a function of vowel openness, backness, and condition (mean \pm 95% confidence interval, adjusted for inter-speaker variability) comparing RVP in the three conditions within vowel groups

The first interaction is due to a difference between the front and back vowel groups in the *V#hV* context (where back vowels exhibited a slightly lower ratio of voicing in /h/, $RVP = 94 \pm 18\%$ for front, and $83 \pm 26\%$ for back vowels), and a somewhat smaller difference observable in the *#hV* context in the opposite direction (where back vowels exhibited a slightly higher ratio of voicing, $RVP = 38 \pm 21\%$ for back and $32 \pm 20\%$ for front vowels). The interactions show that the clearly visible effect of *condition* is to some extent different in the front and back vowel groups (as contexts), while the effect of backness is also to some extent different in the open and close vowels, i.e., neither the back and front vowels, nor the open and close vowels behave as homogenous groups (as contexts) with respect to the RVP in /h/. Despite the interactions, Tukey's post hoc test revealed that all tokens differed in the *#hV* vs. *V#hV* ($p < 0.0001$ in all pairs), and in the *#hV* vs. *VhV* ($p < 0.0001$ in all pairs) comparisons, but not in the *V#hV* vs. *VhV* comparisons, where vocalic features were held constant.

We also analyzed the duration of /h/ as a function of conditions (fixed factor: *condition*) to see if token durations varied systematically in the three conditions. The model revealed that the absolute duration of /h/ is indeed affected by the condition ($F(2, 18.76) = 11.86, p < 0.001$), but according to pairwise comparisons, only the *#hV* condition ($M = 79.89$ ms, $SD = 25.46$) differs from *VhV* ($M = 67.78$ ms, $SD = 18.37$) significantly ($p = 0.006$), while *V#hV* ($M = 75.70$ ms, $SD = 28.20$) does not differ from any of the conditions. Correlation analysis of token durations and RVP additionally showed that the difference between conditions also lies in the interrelation between /h/ duration and the amount of voicing: while /h/ duration was in strong negative correlation with RVP in both types of intervocalic positions (*V#hV*: $\rho = -0.57, p < 0.0001$; *VhV*: $\rho = -0.41, p < 0.0001$), it showed only a very weak significant negative correlation in *#hV* ($\rho = -0.09, p = 0.02$). This result suggests that the duration of the voice leak in /h/ from the right vowel is almost independent of the total duration of the segment post-pausally, but longer segments tend to have less voicing intervocalically (Figure 4).

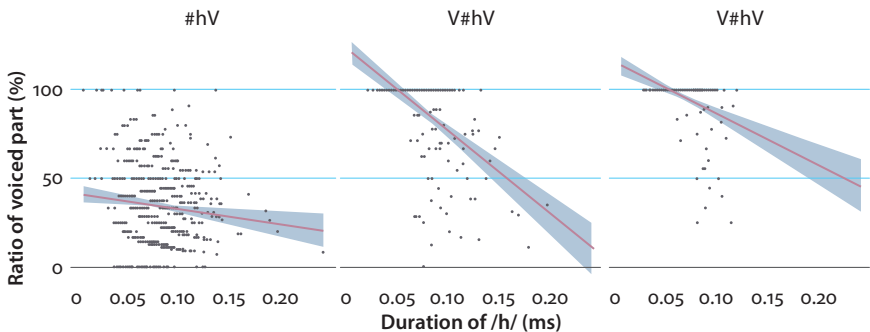


Figure 4. Correlations of segment duration and ratio of voiced part

Lastly, we turn to the measure of harmonicity. Visual observation of the overall HNR data reveal a very similar trend to what we observed in the RVP data, namely that the intervocalic and post-pausal onset conditions differ remarkably, while the two intervocalic conditions pattern together (this observation is also supported by post hoc tests of the linear mixed effects model: for *#hV* vs. *V#hV* $p < 0.029$; for *#hV* vs. *VhV* $p < 0.0001$). /h/ realizations in *#hV* showed dominance of aperiodic (noise) components over periodic components (lower HNR values), while /h/ tokens in *V#hV* and *VhV* appeared to be more periodic (higher HNR values) (Figure 5, upper left). According to the data sorted by the vowel quality feature (Figure 5, upper right), /h/ tokens were the least periodic in /v/ context in all but the *#hV*

condition, while tokens in /i/ context and in V#hV condition were the most periodic. On the basis of HNR values, the following harmonicity order of /h/ realizations emerged in the two intervocalic contexts, in terms of the context vowel: /v/ < /ɛ/ < /u/ < /i/, and the following grouping was established in VhV context: /v/ vs. /ɛ u i/ (where the latter group displayed less noisy /h/ tokens) (Figure 5, upper right).

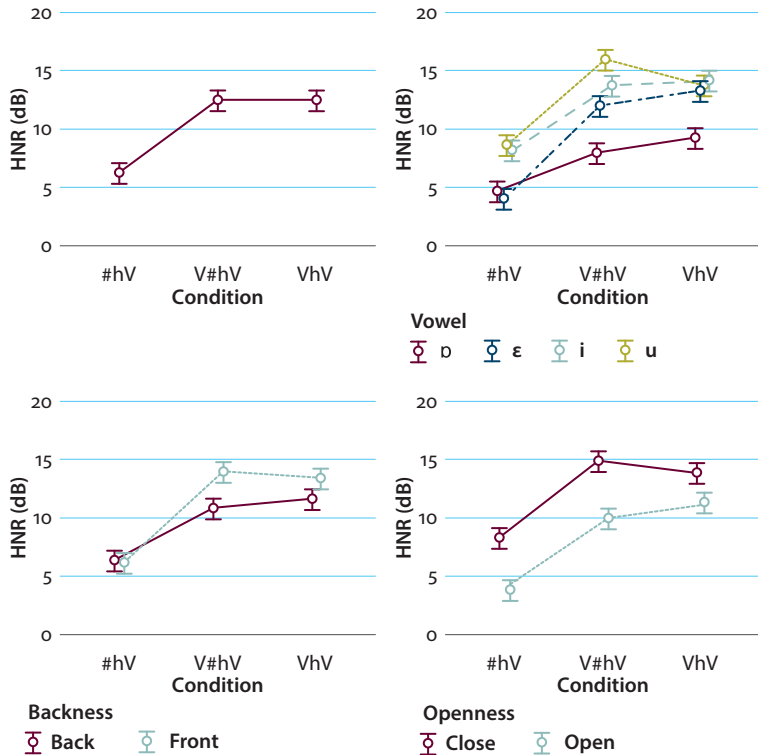


Figure 5. HNR in all /h/ segments pooled (upper left), and as a function of the right vowel's quality (upper right; /v/: black solid line; /ɛ/: black dotted-dashed line, /i/: grey dashed line, /u/: gray dotted line), vowel backness (bottom left), and vowel openness (bottom right) (mean ± 95% confidence interval, adjusted for inter-speaker variability)

On the HNR data, a linear mixed model revealed a significant 3-way interaction effect of *backness*, *openness* and *condition* ($F(2, 995.74) = 13.3670, p < 0.001$), which is to be interpreted as a *condition* by *openness*, or a *condition* by *backness* interaction, that varies across the levels of the third variable. Therefore, in the present design this result is basically equivalent to a *condition* by vowel quality interaction, which reflects again that the open, close, back, and front vowel groups do not behave as homogenous groups with respect to their role in facilitating /h/ voicing.

According to Tukey's post hoc tests, #hV vs. V#hV, and #hV vs. VhV contexts differ in all vowel groups significantly ($p < 0.05$), but the two intervocalic conditions do not. The differences of group means show that the acoustic signal exhibited more dominant periodic components in all conditions in close vowels' context (Figure 5, bottom right), and that a similar trend was true for /h/ tokens in back vowels' context in all but the #hV condition (Figure 5, bottom left).

We also tested if signal harmonicity in /h/ is correlated with RVP, to see to what extent the value of HNR was affected by the length of the periodic signal fragment. We found a strong/moderate positive correlation in both intervocalic contexts (V#hV: $\rho = 0.59$, $p < 0.001$; VhV: $\rho = 0.32$, $p < 0.001$), but no correlation in #hV (Figure 6).

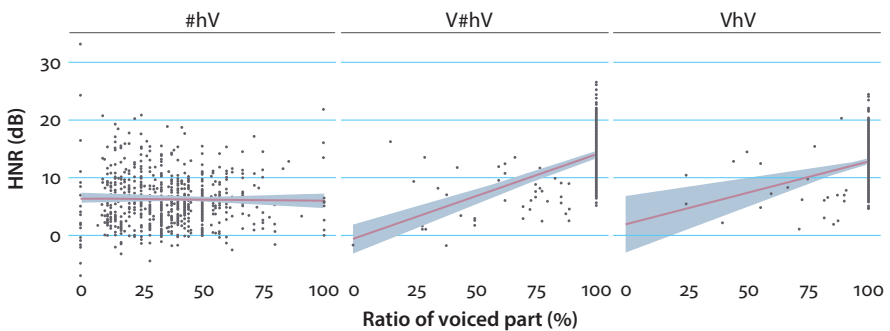


Figure 6. Correlations of HNR and ratio of voiced part

This reflects that the differences we found in HNR values between the intervocalic and post-pausal conditions are linked to the amount of voicing: although an increase in overall signal periodicity is indeed in correspondence with an increased ratio of voiced frames intervocalically, it is not so if the segment is uttered after a pause. In other words, the noisy quality of /h/ tokens in post-pausal position cannot simply be interpreted as a result of shorter relative durations of the voice leak from the neighboring vowel; rather, it is a generally noisier sound quality.

4. Discussion and conclusions

The acoustic analysis of the Hungarian /h/ has shown that the main factor triggering its voicing in syllable onset is the intervocalic position: if the /h/ occurs between two vowels, it is realized as voiced (in roughly 80% of its total duration),

irrespective of the word boundary in the sequence or the presence of pitch-accent on the containing syllable, while in a post-pausal context (regarded as a baseline in the analysis) a very low amount (approx. 30%) of voicing is found. However, vowel backness also has an effect in interaction with this condition. Thus, although in a binary categorization we are safe to categorize /h/ as voiced intervocalically, the exact amount of its voicing is to some extent dependent on the horizontal tongue position of the flanking vowel(s). On a word boundary, and between front vowels, /h/ appears to contain more voicing than between back vowels, but the tendency is reversed when the onset /h/ is uttered after a pause.

The ratio of fully voiced tokens we found in intervocalic positions was highly in line with Mitterer (2018), and roughly in line with the results of Teras (2018). Similarly to Mitterer, we found that 81% of tokens were realized with 100% voicing, while Teras had found this ratio to be 70%. As far as the ratio of voiced variants in intervocalic front and back contexts is concerned, to some extent, our results are close to Gósy (2005), since both studies conclude that front vowels favor /h/ voicing. However, there is a striking difference with regard to the disfavored back vowel context. To recall, Gósy (2005) put the ratio of voiced variants at 75% between front vowels, and 38–39% between back vowels, compared to the present study's findings of 87% between front vowels and 75% between back vowels. We are tempted to put down this divergence in the results to major differences in methodology and speech material applied in the two studies, and conclude that by and large, our results are in accord with Gósy (2005) as well. On the basis of the above, we claim that our first hypothesis is partially corroborated, as the position of the /h/ and the presence of pitch-accent did not show an effect on the amount of voicing in /h/, but the effect of openness is more complex than hypothesized. This effect of vowel quality will receive further comment as we discuss our findings on HNR.

The durational analysis revealed that there are further differences between /h/ variants and the phonetic processes that are responsible for /h/ voicing in the intervocalic and post-pausal positions. The results show that /h/ tokens are the longest if they are produced word-initially and after a pause, and that duration and amount of voicing are in closer correspondence if /h/ is positioned intervocalically rather than post-pausally. While /h/ tends to have less voicing if its total duration increases intervocalically, the amount of voice leak from the neighboring vowel is much more independent of the total duration of the token if it is preceded by a pause.

Our results suggest that the voiced realization of /h/, i.e., [h̥], is voiced in approximately 80% of its total duration, while its voiceless counterpart [h] contains vocal fold vibration in about one third (approx. 30%) of its total duration. As these values were established on the basis of a great number of observations,

they may be also treated as gross reference values in future binary-classification-based analyses.

We also intended to give an initial, exploratory description of the quality of breathy voice Hungarian /h/ is claimed to exhibit. To put it more precisely, we aimed to describe the quality of the sound (but not necessarily just the quality of voice or the voiced fraction) produced during the production of /h/ in intervocalic positions as opposed to the sound quality (may it be voiced, unvoiced, or partially voiced) it exhibits in a baseline post-pausal position. To grasp this general “sound quality” or signal characteristic difference between realizations, we measured and analyzed the harmonics-to-noise ratio as a function of the quality of the flanking vowels, with special attention paid to backness and openness again. We predicted that higher and more fronted positions of the tongue (and the larynx) facilitate the increased adduction of the vocal folds, and thus reduce breathiness, i.e., the amount of noise in /h/. This hypothesis was again partially corroborated by our results, as we found that front vowels and close vowels did indeed facilitate more periodic (less breathy/less noisy) /h/ realizations, but these tendencies were found so clearly only in intervocalic contexts. After a pause, /h/ variants were not distinguished by HNR along the neighboring vowels’ backness feature, while the distinction was still present, only at a generally lower mean value along the openness dimension. HNR data were also in line with RVP (which we used to quantify the amount of voicing), as both of these measures reflected that /h/ tokens are more likely to be voiced (or they contain more prominent periodic components) intervocalically than after a pause. Furthermore, we also showed that the difference we found between the two positions reflected actual dissimilarities of the laryngeal settings the /h/ tokens were produced with, and not only the fact that post-pausal /h/ tokens generally contained a lower ratio of voiced part. In the statistical analysis, however, we found a three-way interaction of the factors, indicating that the vowel quality features openness and backness, which are claimed to distinguish and cluster the tested vowels, do not seem to be relevant, or homogenous categories with respect to their behavior as facilitating contexts for intervocalic /h/ voicing. That is, in this respect, /ɛ/ differs from /v/ in other aspects than /i/ differs from /u/. We can further note that on the basis of the HNR values measured in intervocalic /h/ realizations, two (more or less distinct) vowel groups seemed to emerge: /v/ vs. /i u ɛ/. Before turning to the explanation of this result, we take a short detour into the underlying articulatory mechanisms of these vowels and their traditional featural description (that we also applied in the present study).

Traditionally, vowel qualities are defined and described on the basis of articulatory and perceptual-acoustic features in a mixed manner as follows (see IPA Handbook 1999). The first feature is jaw opening and/or tongue height, i.e., the

vertical position of the highest point of the tongue or the position of the tongue body as a whole. In this dimension we distinguish ‘close’, and ‘open’ vowels, and two in-between categories defined on a perceptual basis: ‘close-mid’, and ‘open-mid’. The second feature is tongue backness, i.e., the horizontal position of the highest point of the tongue or the position of the tongue body as a whole. In this dimension we define ‘front’, ‘back’, and ‘central’ vowels. The third feature is lip rounding. In this dimension we have ‘rounded’, and ‘unrounded’ vowels. (To these features, the additional feature of phonological length is also added, if relevant.) Although these features seem to be rather straightforward, and are thus widely used, if we take into account the actual articulatory maneuvers the vowels are produced by, this ‘lingual oral model’ of classification (Esling 2005: 15) seems to be highly oversimplifying.

In a series of studies done by electromyography and via the computational modelling of the tongue muscles, Honda (1996) analyzed the coordination of the lingual muscles, with special attention paid to vowel articulation. Based on his findings, Honda (1996) proposed that the main four muscles of the tongue are assigned to the production of the anchor vowels of the cardinal vowel system in pairs, as follows: the genioglossus posterior and genioglossus anterior are involved in the production of /i/, the genioglossus anterior and the hyoglossus are employed in the production of /æ/, the genioglossus posterior and the styloglossus are patterning together in the production of /u/, and the hyoglossus and the styloglossus are engaged in the production of /a/. That is, from an articulatory perspective, the difference between “openness” categories is indeed fundamentally different in “front” and in “back” vowels: in “front” vowels (/i/ vs. /æ/) it is indeed the difference in jaw opening, while in “back” vowels (/u/ vs. /a/) it is *raising*, and *retraction* of the tongue in the case of “close” and “open” vowels, respectively (see Esling 2005: 23). What is more, due to the connections of these muscles to other speech organs, these findings also imply that it is the “back open” /v/-like vowels which exhibit the strongest connection to (and the greatest effect on) the laryngeal settings via the direct link of the hyoid bone. In conclusion, it may as well be argued that the coarticulatory effects of /v/-like sounds on the vertical position of the larynx in the flanking speech sounds and thus on the quality of voicing may also be fundamentally different from that of the rest of the “vowel regions” in the vowel space. Therefore, we may now draw the conclusion that our results, which show the two “natural classes” of vowels /v/ vs. /i u ε/ emerging as separate categories, are well-established in the articulatory implementation of the studied vowel categories.

To conclude, the present study has contributed to the observation that the intervocalic voicing of /h/ is a phonetic process that generalizes across languages. It has also shown that such phonological conditioners as pitch-accent do not play

a role in this process, but the openness of the flanking vowels as a vocalic feature does, via coarticulation. Our fine-grained phonetic analysis has further demonstrated that the sound quality of the fricative is also affected by the openness and backness of the flanking vowels. Our proposed explanation for this finding is informed by a less simplified explanation of vowel articulation, and the complex interactions between the tongue and the larynx structures.

Acknowledgements

The authors are very grateful to the two anonymous reviewers for their insightful comments. Help by Gergely Varjasi in gathering experimental data is gratefully acknowledged.

References

- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software Volume* 67(1). 1–48.
- Beckman, Mary. 1995. Implications for phonological theory. In William Hardcastle & John Laver (eds.), *The handbook of phonetic sciences*, 199–225. Cambridge, MA: Blackwell.
- Boersma, Paul. 1993. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *IFA Proceedings* 17. 97–110.
- Boersma, Paul & David Weenink. 2017. *Praat: doing phonetics by computer* [Computer program]. Version 6.0, retrieved 15 July 2017 from <http://www.praat.org/>
- Bolla, Kálmán. 1995. *Magyar fonetikai atlasz. A szegmentális hangszerkezet elemei*. Budapest: Nemzeti Tankönyvkiadó.
- Deme, Andrea, Márton Bartók, Tekla Etelka Grácsi, Tamás Gábor Csapó & Alexandra Markó. 2018. Context dependent voicing characteristics of the Hungarian /h/. *Talk on Hanyang International Symposium on Phonetics and Cognitive Sciences of Language 2018*. Hanyang, South Korea, 18–19 May 2018.
- Demolin, Didier, Sergio Hassid, Thierry Metens & Alain Soquet. 2002. Real-time MRI and articulatory coordination in speech. *Comptes Rendus – Biologies* 325(4). 547–556. [https://doi.org/10.1016/S1631-0691\(02\)01458-0](https://doi.org/10.1016/S1631-0691(02)01458-0)
- Draxler, Christoph & Klaus Jänsch. 2004. SpeechRecorder – A universal platform independent multi-channel audio recording software. In *Proceedings of 4th International Conference on Language Resources and Evaluation*, 559–562. Lisbon. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.58.5793&rep=rep1&type=pdf>
- Eager, Christopher D. 2015. Automated voicing analysis in Praat: Statistically equivalent to manual segmentation. In *Proceedings of ICPHS*. Glasgow, UK: University of Glasgow. <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0083.pdf>
- Esling, John H. 2005. There are no back vowels: The laryngeal articulator model. *Canadian Journal of Linguistics* 50. 13–44.

- G. Kiss, Zoltán. 2013. Measuring acoustic correlates of voicing in stops and fricatives. In Péter Szigetvári (ed.), *VLLXX: Papers Presented to László Varga on his 70th Birthday*, 289–311. Budapest: Tinta Könyvkiadó.
- Garellek, Mark. to appear. The phonetics of voice. In William Katz & Peter Assmann (eds.), *The Routledge Handbook of Phonetics*. http://idiom.ucsd.edu/~mgarellek/files/Garellek_Phonetics_of_Voice_Handbook_final.pdf
- Gobl, Christer & Ailbhe Ní Chasaide. 1995. Voice source variation. In William Hardcastle & John Laver (eds.), *The handbook of phonetic sciences*, 427–461. Cambridge, MA: Blackwell.
- Gósy, Mária. 2004. *Fonetika, a beszéd tudománya*. Budapest: Osiris Kiadó.
- Gósy, Mária. 2005. A /h/ zöngésedése két magánhangzó között. *Beszédkutatás* 5–20.
- Gradoville, Michael Stephen. 2011. Validity in measurements of fricative voicing: Evidence from Argentine Spanish. In Scott M. Alvord (ed.), *Selected Proceedings of the 5th Conference on Laboratory Approaches to Romance Phonology*, 59–74. Somerville, MA: Cascadia Proceedings Project. <http://www.lingref.com/cpp/larp/5/paper2635.pdf>
- Halácsy, Péter, András Kornai, László Németh, András Rung, István Szakadát & Viktor Trón. 2003. A szószablya projekt. In Alexin Zoltán & Dóra Csendes (eds.), *I. Magyar Számítógépes Nyelvészeti Konferencia előadásai (MSZNY 2003)*, 299. http://eprints.sztaki.hu/7886/1/Kornai_1773394_ny.pdf
- Hess, Susan A. 1998. *Pharyngeal articulations*. Ph.D. dissertation. Los Angeles: University of California. http://phonetics.linguistics.ucla.edu/research/Hess_diss.pdf
- Honda, Kiyoshi. 1996. Organization of tongue articulation for vowels. *Journal of Phonetics* 24. 39–52. <https://doi.org/10.1006/jpho.1996.0004>
- Honda, Kiyoshi, Hiroyuki Hirai, Shinobu Masaki & Yasuhiro Shimada. 1999. Role of vertical larynx movement and cervical lordosis in f0 control. *Language and Speech* 42. 401–411. <https://doi.org/10.1177/00238309990420040301>
- Hoole, Philip & Christian Kroos. 1998. Control of larynx height in vowel production. In *Proceedings of the 5th Conference on Language Processing (ICSLP) 2*, 531–534. Sydney. https://www.isca-speech.org/archive/archive_papers/icslp_1998/i98_1097.pdf
- IPA Handbook = International Phonetic Association. 1999. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University.
- Kassai, Ilona. 2005. *Fonetika*. Budapest: Nemzeti Tankönyvkiadó.
- Keating, Patricia A. 1988. Underspecification in phonology. *Phonology* 5/2. 275–292. <https://doi.org/10.1017/S095267570000230X>
- Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13). 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Laver, John. 1994. *Principles of phonetics*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139166621>
- Laziczius Gyula. 1937. A zöngés h kérdése. *Magyar Nyelv* 33. 305–310.
- Laziczius Gyula. 1963/1979. *Fonetika*. Budapest: Nemzeti Tankönyvkiadó.
- Ladefoged, Peter & Keith Johnson. 2011. *A course in phonetics* (6th ed.). Boston: Wadsworth, Cengage Learning.
- Lindblom, Björn. 1990. Explaining phonetic variation: A sketch of the H&H theory. In William Hardcastle & Alain Marchal (eds.), *Speech production and speech modeling*, 403–439. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-009-2037-8_16

- Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511753459>
- Meyer, E. A. & Zoltán Gombocz. 1909. *Zur Phonetik der Ungarischen Sprache*. Upsala: Edvard Berlings Buchdruckerei.
- Mitterer, Holger. 2018. Not all geminates are created equal: Evidence from Maltese glottal consonants. *Journal of Phonetics* 66. 28–66. <https://doi.org/10.1016/j.wocn.2017.09.003>
- Morey, Richard D. 2008. Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology* 4(2). 61–64.
<https://doi.org/10.20982/tqmp.04.2.p061>
- Pabst, Friedemann & Johan Sundberg. 1992. Tracking multi-channel electroglottograph measurement of larynx height in singers. *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)* 33(2–3). 67–78.
- Painter, Colin. 1976. Pitch control and pharynx width in Twi: An electromyographic study. *Phonetica* 33. 334–352. <https://doi.org/10.1159/000259781>
- R Core Team. 2017. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Russell, V. Lenth. 2016. Least-squares means: The R package lsmeans. *Journal of Statistical Software* 69(1). 1–33.
- Siptár, Péter. 1994/2016. A mássalhangzók. In Ferenc Kiefer (ed.), *Strukturális magyar nyelvtan 2. Fonológia*, 183–272. Budapest: Akadémiai Kiadó. Digital edition: 2016.
- Siptár, Péter. 2001. Három felemás magyar mássalhangzó. *Magyar Nyelv* 97. 385–404.
- Siptár, Péter & Miklós Törkenczy. 2007. *The Phonology of Hungarian*. New York: Oxford University Press.
- Sundberg, Johan & Anders Askenfelt. 1981. Larynx height and voice source. A relationship? *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)* 22(2–3). 23–36.
- Stevens, Kenneth N. 1999. *Acoustic phonetics*. Cambridge, MA: MIT Press.
- Szigetvári, Péter 1998. Miért nem zöngésedik a [h]? (draft). In László Büky & Márta Maleczki (eds.), *A mai magyar nyelv leírásának újabb módszerei III*, 169–181. Szeged: JATE. http://seas3.elte.hu/szigetva/papers/miert_h.pdf
- Teras, Pire. 2018. The phonetic variation of short intervocalic /h/ in Estonian. In Katarzyna Klessa, Jolanta Bachan, Agnieszka Wagner, Maciej Karpiński & Daniel Śledziński (eds.), *Proceedings of the 9th International Conference of Speech Prosody*, 883–887. Poznań: Adam Mickiewicz University. https://www.isca-speech.org/archive/SpeechProsody_2018/pdfs/158.pdf
- Zenker, Wolfgang & Adolf Zenker 1960. Über die Regelung der Simmelippenspannung durch von aussen eingreifende Mechanismen. *Folia Phoniatica* 12(1). 1–36.
<https://doi.org/10.1159/000262846>