A PHONETICALLY-BASED APPROACH TO THE PHONOLOGY OF [v] IN HUNGARIAN

ZOLTÁN KISS – ZSUZSANNA BÁRKÁNYI

Research Institute for Linguistics
Hungarian Academy of Sciences
Benczúr utca 33.
H–1068 Budapest
Hungary
{cash|bzsu}@nytud.hu

Abstract: We propose a unified, surface-based functionalist analysis of the phonology of Hungarian *v*, which is shown to fare better than past generative formalist/representational models. The model introduced can account for the two-fold patterning of *v* with respect to voicing assimilation without evoking exceptional means. Furthermore, it can also explain certain asymmetries as well as graduality displayed by *v*'s phonotactic distribution, namely, that some clusters are more frequent in the lexicon, whereas others are marginal. The analysis is grounded in the aerodynamics of *v*'s articulation (which involves inherently contradictory targets) as well as in the relative perceptibility of its contrast in various contexts. It is shown with the help of quantitative experiments that *v*'s phonological patterning is directly derivable from these phonetic factors.

Keywords: phonetics-based phonology, phonotactics, fricatives, voicing, [v]

1. Introduction

The seemingly odd phonological behaviour of [v] in Hungarian has always attracted a lot of attention in the phonological literature. Most attention has focused on its two-fold patterning in voicing assimilation, namely, that it patterns with obstruents in being targeted by the process, but it behaves like sonorants as it does not trigger voicing assimilation. Related to this dynamic aspect, its static phonotactic distribution has also been of interest, as it also displays asymmetrical properties. In this paper, we

provide a functionalist account which can explain [v]'s static and dynamic phonology in a unified manner, based solely on its surface phonetics (in particular, its aerodynamic properties as well as its relative perceptibility in various phonetic contexts).

Having introduced the relevant data in Hungarian (**§2**) and in some other languages (**§4**), as well as the inadequacies of past (formalist/representational) approaches (**§3**), we present the phonetically-based analysis in section **5**. First, the basic functionalist principles that the analysis uses are set forth, then we concentrate on the more specific aerodynamic and perceptual properties of [v], as well as those of the contexts in which it occurs. Based on these basic phonetic principles, we introduce the most important predictions of the analysis concerning [v]'s (i) realizations, (ii) behaviour in voicing assimilation, and (iii) phonotactic patterning. Last, we put forth the results of an acoustic experiment whose primary aim was to characterize the surface realizations of [v] and then check the validity of the proposed predictions, which are shown to be directly related to these realizations.[1]

2. The distribution of [v] and the "Voicing Requirement"

Let us begin with the basic facts concerning [v]'s distribution in Hungarian monomorphemic two-member consonant clusters ([v] does not occur monomorphemically in clusters with more than two consonants in this language). Table 1 below displays the logical possibilities of [v]'s clustering ability in three environments: (i) intervocalically, (ii) word-finally (before a pause), and (iii) word-initially (after a pause). Example words are also provided; in some cases the type frequency of the cluster is low (in other words, it occurs in but a handful of words), such clusters are marked with numbers which represent in how many words they actually

---

[1] In the first portion of the paper, we simply use the IPA symbol [v] to refer to what is usually and traditionally described as the "voiced labiodental fricative". In the second half of the paper, the exact phonetic identity (and variants) of this sound will be made more explicit. Sometimes, however, the orthographic form *v* is used to refer to this consonant in general, without going into details as to its factual realizations. In most cases, we use the orthographic forms of Hungarian words without providing IPA transcription. The IPA transcription of the letters the interpretation of which is non-obvious are as follows: *ty* = [c], *gy* = [ɟ], *sz* = [s], *s* = [ʃ], *zs* = [ʒ], *c* = [ts], *cs* = [tʃ], *dzs* = [dʒ], *ny* = [ɲ]; *a* = [ɒ], *á* = [aː], *e* = [ɛ], *é* = [eː]. An acute accent over vowel letters signals length.

occur. For some clusters, there are no monomorphemic examples, but since they featured in the experiment to be discussed below, they have also been included in the table; these non-monomorphemic clusters are shown bracketed.

Table 1

The distribution of [v] in monomorphemic words (CC clusters)
(based on Siptár–Törkenczy 2000, 98f; 106f; 129f)

	p	t	c	k	b	d	ɟ	g	ts	tʃ/dʒ/f	s	ʃ	v	z	ʒ	m	n	ɲ	l	r	j	h
V_V 1. VvCV						bovden 1		afgán 2											bóvli	sevró 2	szovjet 1	
2. VCvV	lopva	hatvan	kotyvaszt 1	ekvivalens	szubvenció	medve	fegyver 3	dugvány 3			köszvény 3	fösvény		özvegy 2		nyamvadt 4	szenved	ponyva	tolvaj	árva 4	csajvadék 1	
_# 3. VvC#						(hívd)															(hívj)	
4. VCv#						kedv 3										hamv 1	-szenv 2	könyv 2	nyelv 2	szerv	ölyv 1	
#_ 5. #vCV																			vlach 2	Wrangler 1		
6. #CvV		tviszt 1		kvarc		Dvorzsák 1		gvárdián 2	cvekedli 2		szvetter 3	svéd										

The glosses of the words in the table are as follows: *afgán* 'Afghan', *árva* 'orphan', *bovden* 'V-shaped belt', *bóvli* 'trash', *cvekedli* 'pasta with cabbage', *csajvadék* 'vagabond', *dugvány* 'cutting', *Dvorzsák* proper name, *ekvivalens* 'equivalent', *ellenszenv* 'aversion', *fegyver* 'weapon', *fösvény* 'miser', *hívd* 'call.2sg.def.imp.', *hívj* 'call.2sg.indef. imp.', *kedv* 'mood', *kotyvaszt* 'concoct', *könyv* 'book', *köszvény* 'arthritis', *kvarc* 'quartz', *lopva* 'furtively', *medve* 'bear', *nyamvadt* 'lousy', *nyelv* 'language', *ölyv* 'hawk', *özvegy* 'widow', *ponyva* 'canvas', *sevró* 'kid(skin)', *svéd* 'Swedish', *szenved* 'suffer', *szerv* 'organ', *szovjet* 'Soviet', *szubvenció* 'subsidy', *szvetter* 'cardigan', *tolvaj* 'thief', *tviszt* 'twist', *vlach* 'Vlachian', *Wrangler* 'Wrangler jeans'.

It is of course the blank cells of this table that constitute the cases of most interest, that is, the clusters that are missing in the language, as well as those whose type frequency is low. Clearly, the distribution of a single [v] is not restricted intervocalically. The moment the position on either the left- or the right-hand side is occupied by a consonant, distributional restrictions occur, with more restrictions cropping up pre-consonantally, as displayed by the low frequency numbers. Similar observations can be made with respect to the word-final as well as the word-initial position: [v]'s distribution is limited in the context of an adjacent consonant. These distributional effects are summed up in Table 2:

Table 2

The effect of the immediate environment on the distribution of [v]

left env.	[v]	right env.	example	effect on [v]'s distr.
V	[v]	V	*kavics*	no restrictions
C	[v]	V	*medve*	few restrictions
V	[v]	C	*bóvli*	restricted
V	[v]	#	*sav*	no restrictions
C	[v]	#	*kedv*	restricted
#	[v]	V	*vas*	no restrictions
#	[v]	C	[v]*rangler*	restricted

Glosses: *kavics* 'pebble', *medve* 'bear', *bóvli* 'trash', *sav* 'acid', *kedv* 'mood', *vas* 'iron', *Wrangler* 'Wrangler jeans'

Based on Table 2, we can set up a hierarchy of environments, which illustrates how the distribution of [v] is curtailed in various contexts; this is shown in (1), where "$X < Y$" means that Y is an environment where [v]'s distribution is more restricted than in environment X.

(1)   {V——V, #——V, V——#} < C——V < {V——C, C——#, #——C}

Table 1 also (partly) illustrates one of the most salient properties of the phonology of Hungarian obstruents, what we may refer to as the **Voicing Requirement**. According to this requirement, two obstruents standing next to each other may not differ in voicing, that is, either (i) both are voiceless, or (ii) both are voiced (hence the lack of *[vt] or *[vk] clusters). This requirement embraces the whole of Hungarian obstruent phonology; that is to say, it applies morpheme-internally as well as over morpheme and word boundaries.[2] We are going to use the term "Voicing Requirement" (henceforth, "VR") to therefore cover both the "static"

---

[2] Vago (1980, 143) has proven to be wrong when he states that in Hungarian, voice assimilation is optional. It is only when flanking a relatively long pause that two neighbouring obstruents may have different voicing (see Szigetvári 1998a, 223 and Siptár–Törkenczy 2000, 198, who state that "[voicing assimilation] is postlexical (it applies across any type of boundary as long as no pause intervenes) but obligatory and non-rate-dependent"). Unfortunately, Vago's assertion caused other analysts to work with data with empirical inadequacies and consequently reach erroneous conclusions, such as Lombardi (1995; 1999); see, for example, Törkenczy (2000) as well as Kenstowicz et al. (2003) on Lombardi's problematic claims. It must also be noted that Vago does not call voicing assimilation optional in the

and "dynamic" aspects of the phenomenon, what the traditional literature calls **voicing agreement** morpheme-internally and **voicing assimilation** over morpheme and word boundaries.

Word-initial CC clusters constitute a sub-case, where VR is stricter: in this position, obstruent clusters in Hungarian are always voiceless.[3] If we consider [v] a voiced fricative (as the traditional approach does), then this segment is the only regular exception to this generalization, as we do find voiceless/voiced obstruent plus [v] clusters in this position:

(2)   Word-initial obstruent plus [v] clusters

   tviszt 'twist', *kvarc* 'quartz', *szvit* ([sv]) 'suit', *svéd* ([ʃv]) 'Swedish', *Dvorzsák* (proper name), *gvárdián* 'guardian', etc.

In this position thus, [v] patterns with sonorant consonants, which are free to occur here (see (4)).

The "dynamic" aspect of VR works the following way: if two obstruent segments with different values for voicing come to stand next to each other, it is always the second segment that determines the voicing of the first, thus it acts as the trigger of the regressive voicing assimilation. Clusters flanking a boundary between (i) a stem + suffix and (ii) two words (compound words as well as phrases) are affected (so long as no pause obtrudes). The phenomenon is iterative, that is, it can apply to its own output. (3) illustrates this with a few examples:

(3)   Regressive voicing assimilation in Hungarian

   (a)   voicedness assimilated:

   [t][b] → [db]: e.g., *hát-ba* 'back-illat.'; *két#barát* 'two friends'
   [ʃ][b] → [ʒb]: e.g., *has-ba* 'stomach-illat.'; *hús#bolt* 'meat shop'

   (b)   voicelessness assimilated:

   [b][t] → [pt]: e.g., *láb-tól* 'foot-abl.'; *láb#torna* 'foot exercise'
   [z][t] → [st]: e.g., *víz-től* 'water-abl.'; *víz#torony* 'water tower'

   (c)   voicing assimilation is right-to-left iterative:

   [sk][b] → [zgb]: e.g., *groteszk-ben* 'grotesque-iness.'
   [gd][t] → [ktt]: e.g., *smaragd-tól* 'emerald-abl.'

detailed discussion of the rule (pp. 34f), only when he lists it among other rules in the Appendix.

[3] Except for, perhaps, one word *dzéta* 'dzeta' (ʣ), if [dz] is considered a cluster, but its status is not uncontroversial; see, for instance Siptár–Törkenczy (2000, 87ff).

Crucially, VR does not restrict obstruent–sonorant/sonorant–obstruent clusters. In monomorphemic words, both voiced as well as voiceless obstruents can stand next to a sonorant. Also, a sonorant will not voice a preceding voiceless obstruent and a voiceless obstruent will not devoice a preceding sonorant.[4]

(4) Obstruent–sonorant sequences (monomorphemic)

plakát 'poster', *próba* 'rehearsal', *tréfa* 'joke', *knédli* 'dumpling', *klarinét* 'clarinet', *krém* 'cream', *gnóm* 'gnome', *grafika* 'graphics', *friss* 'fresh', *szmog* 'smog', etc.

(5) Lack of regressive voicing assimilation in obstruent–sonorant sequences

    (a) Sonorants are not devoiced:

        [m][t] →[mt] (*[m̥t]): *rém-től* 'monster-abl.'

        [l][t] →[lt] (*[l̥t]): *hal-tól* 'fish-abl.'

    (b) Sonorants do not voice:

        [p][n] →[pn] (*[bn]): *kép-nél* 'picture-adess.'

        [s][n] →[sn] (*[zn]): *rész-nél* 'part-adess.'

Turning to [v], we can state that it—apparently—behaves asymmetrically with respect to VR: it undergoes devoicing (6a),[5] but does not trigger voicing (6b):

(6) [v] and the Voicing Requirement

    (a) [v][t] →[ft]: *sav-tól* 'acid-abl.'

        [v][h] →[fh]: *sav-hoz* 'acid-all.'

    (b) [t][v] →[tv] (*[dv]): *két vár* 'two castles'

        [p][v] →[pv] (*[bv]): *szép vár* 'nice castle'

Thus pre-obstruent ("coda")/target [v] behaves as an obstruent, while post-obstruent (and prevocalic) ("onset")/trigger [v] patterns with sonorants.

---

[4] This is true of the standard dialect of Hungarian, or "Educated Colloquial Hungarian (ECH)" (see Siptár–Törkenczy 2000, 3), the dialect of Hungarian which this paper discusses. However, there are dialects (in Western Transdanubia) where anterior nasal sonorants trigger voice assimilation: *kismiska* [kiʒmiʃkɒ] 'trifle', *hát nem* [haːdnɛm] 'well not', etc. (see Fodor 2003, 339). Interestingly, and—as we will see (Table 4)—perhaps not accidentally, it is these dialects that display voice assimilation triggered by [v], too, as well as word-final devoicing.

[5] As we have seen above, this aspect of [v]'s behaviour is statically satisfied by the fact that there are no monomorphemic words with a [v] plus voiceless obstruent cluster.

A further generalization, the last one here, that we can draw from Table 1 is the following: in word-final position, [v] patterns with obstruents as it can cluster with sonorants as the second consonant. Sonorants do not normally occupy such a position.[6]

(7)   Word-final C[v] clusters (complete list)

[mv]:   *hamv* 'ash' (normally occurs suffixed as in *hamvai* 'his/her ashes')
[nv]:   *ellenszenv* 'aversion', *rokonszenv* 'sympathy'
[ɲv]:   *könyv* 'book', *enyv* 'glue'
[lv]:   *elv* 'principle', *nyelv* 'language'
[rv]:   *terv* 'plan', *szerv* 'organ', *érv* 'argument', *konzerv* 'tinned food', *ismérv* 'criterion', *keserv* 'sorrow', *mérv* 'extent', *orv* 'vile', *örv* 'guise', *sérv* 'hernia', *szarv* 'horn'
[jv]:   *ölyv* 'hawk'
[dv]:   *kedv* 'mood', *nedv* 'fluid', *üdv* 'salvation'

It is true though that there are not many examples of sonorant plus *voiced* fricative clusters word-finally anyway, some of the examples that the literature cites are obsolete as free forms (e.g., *nemz* 'beget', *tömzs* ([mʒ]) 'lode'). The most frequent examples involve [r] (e.g., *borz* 'badger', *törzs* ([rʒ]) 'trunk', etc.); this is true of [v], too: [v] is rare after sonorants other than [r] in word-final clusters.

Note also the three examples in the last row of (7): they involve [d] as the first member of the cluster. This cluster is also rare (and exceptional if one considers the sonority sequencing violation), but as Siptár and Törkenczy (2000, 80) put it, "the [sonority sequencing] violation is at least not unprecedented [(see, e.g., *edz* 'train_V', *pedz* 'begin to understand')] if /v/ is a fricative (an obstruent), whereas if it is a sonorant, [words like *kedv* 'mood'] would violate the otherwise exceptionless generalization that (on the surface) no final cluster can consist of a sequence of obstruent plus sonorant."

The generalizations regarding [v]'s behaviour that we detailed above can thus be summed up as follows: prevocalic [v] in a syllable onset behaves as a sonorant, while a [v] syllabified in a coda patterns and thus behaves as an obstruent. The two types of behaviour are manifest in [v]'s distribution in CC clusters as well as its patterning with respect to the Voicing Requirement:

---

[6] Two exceptions include the coronal liquid [l], which marginally occurs after [r] and [j] in recent borrowings and names: *görl* 'chorus girl', *fájl* 'file', *geil* 'yucky'. [j] does occur after sonorants word-finally; however—and crucially, as we will see later on—, its phonetic manifestation is a noisy fricative [j] (which often devoices to [j̊]): *férj* 'husband', *szomj* 'thirst'.

(8)   Sonorant behaviour of [v]

    (a)   phonotactics: can stand with obstruents word-initially, like the other sonorants:
        tviszt 'twist', *kvarc* 'quartz' ∼ *tréfa* 'joke', *klarinét* 'clarinet'

    (b)   VR: "trigger" (post-obstruent/prevocalic) [v] does not satisfy VR (even though it is voiced), like the other sonorants:
        hatvan 'sixty' *[dv] ∼ *hátra* 'backwards' *[dr], *paplan* 'duvet' *[bl]

(9)   Obstruent behaviour of [v]

    (a)   phonotactics: can stand after sonorants word-finally, like the other obstruents:
        könyv 'book', *terv* 'plan' ∼ *vonz* 'attract', *torz* 'distorted'

    (b)   VR: "target" (pre-obstruent) [v] satisfies VR, like the other obstruents:
        [v][t] → [ft]: *sav-tól* 'acid-abl.' ∼ [z][t] → [st]: *láz-tól* 'fever-abl.'

Here we must note that in the Western Transdanubian ("Nyugat-Dunántúl") dialect of Hungarian, "sonorant" [v] can trigger voicing assimilation; in this variety, other sonorants (most notably anterior nasals—see footnote 4) can also act as voicing triggers, and so [v] patterns with sonorants even here. Examples are listed in (10a) ("ECH" stands for Educated Colloquial Hungarian):

(10) [v]'s voicing behaviour in the Western Transdanubian dialect
      of Hungarian (WTH) (data from Imre 1971 and Juhász 2003)

    (a)   [v] voices a previous voiceless obstruent:
        borotva 'razor': WTH [dv] (ECH [tv])
        rakva 'put.part': WTH [gv] (ECH [kv])
        vasvilla 'iron fork': WTH [ʒv] (ECH [ʃv])

    (b)   [v] devoices after a voiceless obstruent:
        ötven 'fifty': WTH [tf] (ECH [tv])
        csukva 'closed': WTH [kf] (ECH [kv])
        húsvét 'Easter': WTH [ʃf] (ECH [ʃv])

In (10b), we can see another strategy in voiceless obstruent–[v] clusters, namely the progressive assimilation of voicelessness. This latter strategy is common in Vas county and the northern areas of Zala county. This type of assimilation is said not to apply over word boundaries (including compound words). According to Imre (1971, 263) and Fodor (2003, 339), both strategies are noticeable in three areas of Western Transdanubia: Zala county, the southern regions of Hetés and in Baranya county, even in the speech of the same speaker.

Works describing the "Janus-faced" behaviour of [v] often draw a parallel between its two-fold patterning and its phonetic manifestation. For example, Siptár notes that "[[v]'s phonetic realizations form a **continuous range** of smooth transitions from a thoroughbred **fricative** with a lot of noise of friction to weak and noiseless **approximant**" (1996, 83; emphasis ours). Also, as Siptár (*op.cit.*, 88) as well as Siptár–Törkenczy (2000, 80, footnote 7) state, "phonetically, the degree of friction seems to correspond nicely to the [two-fold] pattern [displayed by [v]]." Specifically, when [v] is realized very noisy, its behaviour patterns with that of obstruents; when it is realized as an approximant, it displays a sonorant-like behaviour. Both works list examples that are meant to illustrate this correspondence between behaviour and phonetic realization. Table 3 below is an attempt at interpreting these claims concerning [v]'s phonetic realizations.

Table 3

The allophonic realizations of [v]

(after Siptár 1996, 83 and Siptár–Törkenczy 2000, 80)[7]

most fricatival:	1.	Dv#	*terv* 'plan'	"obstruent"
		vT	*hívsz* 'you call', *óvtam* 'I protected'	
	2.	VvDV	*révbe* 'to port', *bóvli* 'junk'	
		Vv#	*sav* 'acid'	
	3.	VDvV	*medve* 'bear', *olvas* 'read'	"sonorant"
		VvV	*kova* 'flint'	
		#vV	*vér* 'blood'	
least fricatival:	4.	#CvV	*kvarc* 'quartz'	
		VTvV	*pitvar* 'porch'	

According to the table above, the allophones of [v] can be grouped into four sets with respect to their fricativalness: the members of group 1 display the most friction, those of 4 the least. Group 2 and 3 occupy an intermediate place between these two extremes. The double line separates the [v]'s according to their phonological behaviour: groups 1–2 behave like obstruents, whereas groups 3–4 as sonorants. Notice that the [v] in *terv* 'plan' is claimed by Siptár (1996) and Siptár–Törkenczy (2000) to have the same degree of friction as the assimilated (fully devoiced) [v] (usually alleged to be a [f]) of *hívsz* 'you call' and *óvtam* 'I protected'.

---

[7] D = voiced consonant (including sonorants), T = voiceless obstruent. "#" is meant to signal absolute word-final/-initial position (a pause).

It is not made explicit at all whether the [v] in *terv* (and *kedv, könyv, szenv* etc., see (7)) is also devoiced or not. Actually, what one can infer from the works cited is that it preserves its voicing—and also maintains its friction, because Hungarian is claimed to display **no word-final obstruent devoicing**.[8] This aspect of [v]'s phonetics (voicing vs. friction) will be crucial in the phonetically-based analysis to be presented here. A major problem of this classification is that it is only based on the intuition of the authors, it is not backed up by any experimental evidence whatsoever. The bulk of the paper will thus be devoted to testing these claims regarding [v]'s realization. It is, however, not difficult to see that there does seem to be a strong link between [v]'s phonological behaviour in Hungarian and its phonetics; the problem is that no theory has been able to make this link explicit and explanatory. The approach of this paper will try to prove that the link is valid and can be used to explain the seemingly odd behavior of [v].

At this point it is appropriate to mention two apparent discrepancies between [v]'s patterning and its phonetic realizations. As both Siptár (1996, 88) and Siptár–Törkenczy (2000, 203) mention, in post-obstruent position, when the obstruent in question is a labial stop ([p, b]),[9] [v] is said to be a rather strong/noisy fricative, yet it behaves phonologically as a sonorant because it does not trigger voicing assimilation on the previous [p]:

(11) V[pv]V and V[bv]V clusters

[pv]: *lopva* 'stealthily', *alapvető* 'basic', *képviselő* 'representative', etc. (*[bv])
[bv]: *dobva* 'throwing', *szabvány* 'standard', *szubvenció* 'subsidy', etc.

The other minor hitch concerns word-initial #[v]C clusters. Examples for such clusters are not easy to come by; nevertheless, they include *Wrangler* 'a pair of Wrangler jeans', *vlach* 'Vlachian' (also occurs as *Vlach*, personal name) and *Vladimir* (personal name). As both Siptár (1996) and Siptár–Törkenczy (2000) claim, the [v] in these words is a fricative phonetically, yet it again behaves as a sonorant since it does not voice a previous voiceless obstruent:

(12) *két Wrangler* 'two pairs of Wrangler jeans': [tvr] (*[dvr])
két vlach 'two Vlachians': [tvl] (*[dvl])

---

[8] This is true of ECH; see footnote 4, however.

[9] Petrova–Szentgyörgyi (2004) actually extend the environment to include **all** labials, thus also including [m].

Interestingly, this behaviour is mirrored in similar clusters, as in #*hr*-[xr]: *Hradzsin* 'Hradzin (castle in Prague)', *Hruscsov* 'Khruschev' are articulated with a fricatival [x], even though in non-cluster onsets *h* is realized as the glottal glide [h] (single [x] occurs in coda position: *doh* [dox] 'musty smell', *jacht* [jɒxt] 'jacht').[10] This fact is mentioned in Siptár (1996) and Siptár–Törkenczy (2000), but no formal link is made between the phonetic fact and the phonology (allophony) of these segments despite the apparent parallel.

To sum up, the most important questions that this paper seeks to answer are the following:

(13) (a)   What are the factors that direct [v]'s **phonotactic patterning**?
            (see Tables 1 and 2, (8a), (9a))

     (b)   What are the factors that lie behind [v]'s behaviour with respect to the **Voicing Requirement**? (see (8b) and (9b))

     (c)   Are [v]'s phonotactics and its behaviour with respect to the Voicing Requirement **related**? If they are, what is the reason for that?

     (d)   What **acoustic phonetic features** characterize the various realizations of [v]?

     (e)   Are [v]'s phonotactics and its behaviour with respect to the Voicing Requirement related to its **phonetic characteristics**? If they are, in what ways and how can it be modelled?

3. Previous approaches to the phonology of Hungarian [v]

The apparently asymmetrical behaviour of Hungarian [v] has attracted a lot of attention in the phonological literature over the past decades. In this section, we will provide a brief overview of these past approaches. The common denominator in them is the use of the generative **formalist/representational** (thus non-functionalist) model. (14) enumerates the most common suggestions that have emerged:

(14) (a)   positing two underlying segments (obstruent [v] as well as sonorant [v])

     (b)   sonorant [v] is underlying, obstruent [v] is derived
            (Szépe 1968; Vago 1980; Olsson 1992; Siptár 1994)

     (c)   obstruent [v] is underlying, sonorant [v] is derived

     (d)   underlying [v] has an intermediate sonority value
            (Barkaï–Horvath 1978)

---

[10] A detailed description (and an OT analysis) on the allophony of Hungarian *h* can be found in Siptár–Szentgyörgyi (2002) and Szentgyörgyi–Siptár (2005).

(e)   the use of an additional binary feature [±transient] (Zsigri 1994)

(f)   feature geometrical approach (Kornai 1994)

(g)   the use of underspecification/syllable constituents (Siptár 1996; Siptár–Törkenczy 2000)

(h)   non-derivational approaches: Government Phonology (Szigetvári 1998a;b; Cyran 1997; Cyran–Nilsson 1998); Head-Driven Phonology (Ritter 2000); Optimality Theory (Petrova–Szentgyörgyi 2004; Blaho 2005)

In this paper, we cannot offer a comprehensive review of all these proposals due to space restrictions; here we only focus on the most important aspects of them.[11]

The claim that the double-faced behaviour of [v] can be explained by assuming two distinct underlying segments is usually dismissed in the classical phonemic and generative literature very quickly on the following grounds. First, the distribution of either segment is predictable: they are in complementary distribution. Also, the two allophones ("obstruent [v]" and "sonorant [v]") are phonetically similar. This then constitutes a classic case for allophony, with **one** allophone to be abstracted away as the underlying phoneme of the allophonic alternation. Furthermore, the alternation is absolutely automatic: a word-final [v] (claimed to be phonetically a voiced fricative [v]) is realized as an approximant sonorant (often transcribed as [ʋ]) when, for example, a vowel follows it: *sav* [ʃɒv] ∼ *savas* [ʃɒʋɒʃ] 'acid' ∼ 'acidy'.

If we follow the reasoning of the previous paragraph, then one choice available is to assume [v] to be a **sonorant** underlyingly. One of the earliest classical generative accounts—and without doubt the most influential one—couched in these terms is Vago's (1980). In his binary distinctive feature system, this consonant is thus [+ son]. The rewrite rule that Vago (*op.cit.*, 35) postulates to account for voicing assimilation is given in (15):

(15) Voicing Assimilation *à la* Vago (1980) (1)

$$[- \text{son}] \rightarrow [\alpha \text{voi}] \Big/ \underline{} (\#) \begin{bmatrix} - \text{son} \\ \alpha \text{voi} \end{bmatrix}$$

By assuming [v] to be [+ son], underlying [v]'s will thus not be affected by Voice Assimilation. This way Vago is successful at accounting for those [v]'s that behave like sonorants (which do not voice a previous obstruent). The problem, of course, concerns the ones that behave like obstruents:

---

[11] A more thorough review can be found in Kiss (in preparation).

those that devoice before a voiceless obstruent. To achieve this, he needs to introduce an additional voicing assimilation rule which only refers to [v]. The two rules can be collapsed, as shown in (16):

(16) Voicing Assimilation *à la* Vago (1980) (2)

$$
\left\{
\begin{array}{l}
[-\text{son}] \\
\left[
\begin{array}{l}
+\text{cons} \\
-\text{cor} \\
+\text{cont}
\end{array}
\right]
\end{array}
\right\}
\rightarrow [\alpha\,\text{voi}] \bigg/ \underline{\hspace{1cm}} (\#)
\left[
\begin{array}{l}
-\text{son} \\
\alpha\,\text{voi}
\end{array}
\right]
$$

(Obstruents and [v] are assimilated to a following obstruent in voicing.)

Furthermore, Vago (*op.cit.*, 46, note 6) assumes a "low-level adjustment rule" that makes [f] derived from [v] by rule (16) switch from [+son] to [−son], a kind of redundancy rule according to which, all voiceless segments are automatically [−son], too.[12]

The stance taken by Vago (1980) embraces many subsequent approaches. Therefore, the inadequacies that Vago's analysis suffers from are also transferred to those models. In addition to the problematic aspects of the theoretical foundations (namely, those of the derivational, rewrite rule-based system), the analyses are also handicapped from being unable to provide a unified explanatory analysis of [v]. The formalist models of the past are usually capable of accounting for only **one** aspect of the two-faced patterning of this consonant. If, for instance, [v] is assumed to be a sonorant underlyingly (as is the case in Vago 1980), only its sonorant face is explained; so that its obstruent behaviour may be accounted for, arbitrary, stipulative and often exceptional measures are taken. They come in the shape of late "adjustment rules", radical feature changes, exceptional feature geometrical configurations, absolute neutralization rules, redundancy rules, which are posited to operate at the "systematic phonetic level" in the "phonetic interpretation module".[13] Heavy

---

[12] By assuming [v] to be a sonorant, Vago (1980) is also successful at capturing the static phonotactic aspect of [v]'s sonorant behaviour, viz., it can stand after obstruents in initial clusters (8a). However, a separate rule is still required to account for its static obstruent behaviour (9a). (Vago himself does not mention the static aspects of [v]'s phonology.)

[13] Olsson (1992), with his unorthodox "as if rules" (sonorant [v] behaves "as if" it were an obstruent [v] before a consonant and in word-final position), as well as Szépe (1968) using abstract [β] to underlie [v]'s surface allophones are perhaps the most typical examples of the unconstrainedness of generative derivational analyses.

use is made of this module (see especially Siptár's (1996) paper), but no principled analysis is given concerning this compartment of the grammar in any of the works that employ it.

Some works try to avoid the pitfalls of having to classify [v] either as a sonorant or an obstruent by assuming it to rather occupy an intermediate place. The earliest proposal along these lines is Barkaï–Horvath (1978). Their analysis draws on the idea of the **sonority hierarchy**; they propose the following sonority scale of segment classes and individual segments:

(17) Sonority hierarchy (Barkaï–Horvath 1978, 83)

stops	fricatives	[v]	nasals	[y][14]	[r]	[l]
1	2	3	4	5	6	7

Accordingly, the feature [sonorant] will have seven values, and the voicing assimilation rule (which they posit to be active in Hungarian, Hebrew and Russian) will refer to these values:

(18) Voicing Assimilation *à la* Barkaï–Horvath (1978)

$$[m\,\text{sonorant}] \rightarrow [\alpha\,\text{voice}] \Big/ \underline{\hspace{2cm}} \begin{bmatrix} n\,\text{sonorant} \\ \alpha\,\text{voice} \end{bmatrix}$$

where $m \leq 3$ and $n \leq 2$

That is, a segment with less than and including 3 on the sonority hierarchy scale undergoes voicing assimilation before any class which does not exceed 2 on this scale.

Certainly, this solution directly captures the intermediate attribute of [v] with respect to voicing assimilation ([v] undergoes the rule, but does not trigger it), but its phonetic reality is questionable (namely that [v] **always** has an absolute sonority value of "3", regardless of its position, context), and it is difficult to see how a multivalent feature system like this can be managed in a principled/constrained way (e.g., why is it only sonorancy that is multivalent?, how exactly does the phonological module operate on non-binary feature values?, etc.). Also, the way *SPE* (Chomky –Halle 1968) defines [sonorant] and the meaning of sonority does not seem to be fully compatible. And lastly, the formalism does not capture the phonological behaviour of [v] in an explanatory way.

---

[14] Barkaï and Horvath's [y] stands for IPA [j].

Another attempt to reflect on the intermediacy of [v] can be found in Zsigri (1994). He actually introduces a new feature [±transient], which divides the various segment classes, and [v], in the following way:

(19)

	obstruents	[v]	nasals	liquids/glides
[sonorant]	−	−	+	+
[transient]	−	+	−	+

This new feature is to be interpreted as follows: "[+ transient] is present in consonants during the articulation of which there is neither total closure nor fricative noise" (Zsigri *op.cit.*, 19). Since this basically defines non-nasal sonorants, [nasal] is discarded in Zsigri's system. Zsigri's (1994) way of representing voice assimilation is shown in (20):

(20) Voicing assimilation *à la* Zsigri (1994, 69)

$$
\text{ROOT} \quad [- \text{son}] \qquad \begin{bmatrix} - \text{son} \\ - \text{trans} \end{bmatrix}
$$
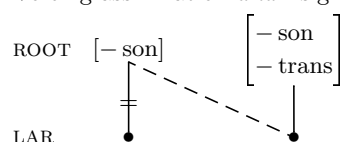
This way, voice assimilation can be represented by using a single rule and [v]'s behaviour with respect to it is captured correctly. It is, however, difficult to see how this new feature (and specifying [v] as [− son]) fares better than using the combination [±son, ± nasal], which divides the segments in (19) the same way, **if [v] is specified as [+ son]**. A step like that would of course take us back to square one, as it were, to the proposal of Vago (1980).[15] According to Zsigri (*op.cit.*, 22), [±transient] is also useful in expressing the phonotactic generalization concerning obstruent [v] as the second segment in "complex onsets". This is necessary in his framework as [v] is an obstruent, and so if it was rather classified as a sonorant, [±transient] would not be necessary, and with [±sonorant], the

[15] Zsigri (*op.cit.*, 67, 69) also claims that his system with the feature [± transient] and [− son] [v] is better at formulating the behaviour of [v] in Western Transdanubian voicing assimilation (see (10)): the rule in those dialects would only contain a triggering [− son], which will thus include [v]. The analyses where [v] is [+ son] would need to have three rules to express the dialectal assimilation. He also notes that positing different underlying [v]'s in the triggering vs. non-triggering dialects would be infelicitous, as there is no difference on the surface between the [v]'s. It is not clear why this is such a great problem, considering the capabilities of the derivational model (e.g., the use of late, "adjustment" rules), which are made great use of in his own analysis anyway.

relevant well-formedness condition could safely be expressed (as is done in Törkenczy 1994, see below).[16] Lastly, in addition to its dubious phonetic definition, the use of [± transient] cannot escape the difficulties of phonetic implementation: Zsigri, too, needs to resort to various default rules to ensure that a [f] devoiced from a [v] does not remain [+ transient] and a [v] voiced from a [f] becomes [+ transient].[17]

Even the frameworks that base themselves on more constrained representational foundations (such as autosegmental phonology, underspecification theory, the non-derivational frameworks: Government Phonology and Head-Driven Phonology (with their element theory and the giving up of the systematic phonetic level/module) as well as Optimality Theory) cannot fully cope with the behaviour of [v]. It actually turns out that in some cases (most notably in Government Phonology), the framework itself is too constrained to be able to offer a comprehensive account of the facts. As shown in Szigetvári (1998a;b), GP-based models face serious problems when explaining how an underlyingly sonorant [v] becomes an obstruent: if a strengthening process like this is assumed to involve the gaining of phonological elements, it is not clear at all how elements are acquired, from "nowhere", as it were. Furthermore, the fortition (obstruentization) of [v] is to take place in an environment (word-finally and pre-consonantally) where normally lenition phenomena occur.[18]

The problem of which variant of [v] is to be posited underlyingly apparently disappears in Optimality Theoretic analyses, owing to the OT architecture and its principle of "Richness of the Base", which states that there are no constraints on inputs; hence either [v], or [ʋ], or, an underspecified [v] can all be assumed underlyingly, and the grammar will still select the grammatical output. This stance is put forth in Petrova–

---

[16] [v]'s phonotactics in C__# (especially in *kedv*-like words, see (7)) remains to be a problem for the [± transient]-based analysis, but it has always been a problem for **any** formalist approach anyway.

[17] One more, independent, motivation for the use of [± transient] is its role in the vocalisation of transient consonants: transient sounds are hypothesized to be "capable of an easy alternation with vowels"; for details, see Zsigri (1994, 19).

[18] In Cyran–Nilsson (1998) [v]-obstruentization is represented as a change in the intrasegmental head–dependency relations ([w] { U } → [v] { U }), and so the problem of fortition not having a local source is circumvented. An alteration like that is in itself problematic (because of the change in the governing/licensing relations: a headless element is promoted for head status), but it still does not explain why it happens in a typical **lenition** site; furthermore, in their system, the voicing assimilation of [v] → [f] would still have to involve elements (such as "noise") appearing from nowhere.

Szentgyörgyi (2004), a paper focusing on [v] in both Hungarian and Russian, two languages where this consonant displays very similar behaviour. The paper starts off by setting up the constraints and their hierarchy, which can account for the voicing assimilation effects. If only these constraints were to be used for clusters with [v], then their system would generate ungrammatical outputs. This happens when [v] is prevocalic:[19] in the case of *ötven* [øtʋɛn] 'fifty' the (wrong) output *[ødvɛn] is selected as optimal; the same happens in the case of *svéd* [ʃʋeːd] 'Swedish', where *[ʒveːd] is selected as optimal. To remedy this situation, they introduce a constraint which basically directs the conditions under which [v] can surface as a sonorant. This constraint, together with another one, which does not allow sonorants to voice, will now work well in selecting the actual grammatical candidates as optimal.[20] The two constraints referred to are shown below:

(21) (a) LAB SON
         A voiced labiodental continuant is a sonorant if and only if it is followed by
         a sonorant.

     (b) ID PRESYL SON VOICE
         A segment before a syllabified sonorant in the output and its input corre-
         spondent must have identical specifications for voice.

A specialty of the constraint LAB SON in (21a) is that it does not refer to any prosodic constituent: it is only sensitive to the **linear**, immediate right-hand environment of [v], a rather different approach to other frameworks (among them Blaho's 2005). Additionally, notice that LAB SON particularizes that prevocalic [v] be a sonorant, whose behaviour is then directed by ID PRESYL SON VOICE (which will not let it voice a previous voiceless obstruent). This is clearly an attempt thus to connect [v]'s phonetic property with its phonological patterning with regard to the "Voicing Requirement".

The burdensome cases, the *lopva*-type and the *Wrangler*-type words, are also problematic for Petrova–Szentgyörgyi (2004). Their OT system will wrongly predict sonorant [ʋ] to surface in each of these cases. They acknowledge this issue, and propose the following constraint to resolve it.

---

[19] Note that Petrova–Szentgyörgyi (2004) posit inputs with a [−son] [v] in the following examples, thus, [øtvɛn] for *ötven* 'fifty' and [ʃveːd] for *svéd* 'Swedish'.

[20] We refer the reader to the paper of Petrova–Szentgyörgyi (2004) for the full evaluation analysis as the exact mechanism itself is not the main concern of the present discussion.

(22) *LABIAL + [v]
      Labial consonants are not followed by a labial approximant.

Obviously, this (very specific) constraint must be ranked higher than
LAB SON. Unfortunately, useful as it may seem, this analysis runs into
problems — due to the constraint system already proposed to account
for the other cases. Recall that LAB SON does not allow [v] to occur
before a (syllabified) sonorant; therefore, neither [ʋ] nor [v] is allowed to
surface in, say, *lopva*. Since Petrova and Szentgyörgyi's model also allows
for underspecification, candidates with [v] underspecified for sonorancy
(which they mark as [V]) are also generated, usefully enough, because in
fact such an underspecified candidate is selected as optimal for *lopva* if the
constraint in (22) is employed (and ranked high): *lo*[pV]*a*. So, a crucial
role has to be played by the phonetic interpretation module—once again:
this module will interpret an underspecified [V] as an obstruent by filling
in the missing [− sonorant] feature.
      To account for [v] + sonorant consonant clusters (in which a noisy
[v] is supposed to surface),[21] Petrova and Szentgyörgyi make use of yet
another constraint, *WC:[22]

(23) *WC
      The feature [+ sonorant] may not be associated with a labiodental continuant
      before a (sonorant) consonant.

This highly ranked constraint (above *LABIAL + [v] and LAB SON) will
not allow [ʋ] before sonorants, and so, again, a [V] will be selected as
optimal ([v] + sonorant will be blocked by LAB SON), which phonetic in-
terpretation will realize as a noisy, obstruent [v].
      To sum up, Petrova and Szentgyörgyi's (2004) OT analysis of [v] in
Hungarian, with its array of constraints as well as assumptions regarding
input forms (underspecification), is (by and large) adequate descriptively.
Phonetic factors do seem to be part of the analysis, most of the constraints

---

[21] The authors refer to **all** presonorant *v*'s here; but their claim that a *v* before a
   sonorant consonant is **always** realized as a noisy [v] is empirically wrong; see **§5.3.2**.

[22] *WC is a specialized version (relativized to a following sonorant) of the more
   general constraint *W, which bans voiced labial approximants (thus supposedly
   [ʋ], too). This latter constraint is posited by Petrova–Szentgyörgyi (2004) to
   account for the apparent fact that Russian *v* never surfaces as a sonorant [ʋ], but
   it still behaves in a dubious way phonologically. However, see Padgett (2002) and
   Lulich (2004), who show that Russian *v* does surface as a (narrow) approximant
   in some contexts.

are phonetically driven and claimed to be phonetically motivated, too.[23]
However, phonetic facts are not the direct explanatory drives behind the
phonological patterning in their model. It is therefore only apparently
and partially phonetically-grounded: their crucial employment of under-
specification emphatically shows that in some cases, a phonology-external
compartment is supposed to arrive at a certain phonetic realization, which
is phonologically relevant; a truly phonetically-based system would do
the reverse: the phonetic realization of a segment (or rather, its phonetic
characteristics and those of its linear surroundings) would be motivating
phonological patterning **within the phonology**.[24] This would be especially
fitting when there is a clear-cut linkage between behaviour and realization
(like in the case of Hungarian [v]) in every position it finds itself. Notice
also that Petrova and Szentgyörgyi's model crucially relies on syllable
structure (see the underspecification of unsyllabified sonorants) and so
the claim that in the case of **some** constraints, linearity (the avoidance of
syllabic constituency) is an advantage cannot be accepted without reser-
vations: a more compact (and truly phonetics-based) model would not,
and perhaps should not, need to make reference to prosodic constituents,
this would be **genuine** linearity. Also, [v]'s homogeneous behaviour is ex-
pressed as if it were the result of separate, independent constraints. In
this paper, we will try to show that a single phonetically-based explana-
tion can account for the static/dynamic behaviour of this consonant (and
other, phonetically similar, ones) in the various positions it finds itself.

So far we have only been discussing past approaches to the analy-
sis of [v]'s phonological behaviour in Hungarian voicing assimilation, but
what about the generalizations concerning its phonotactics, those that we
discussed at the beginning of the paper? As long as some of the phono-
tactic constraints concerning [v] can be related to what we have called
the static manifestation of the "Voicing Requirement", then, of course,
the previous analyses did tackle this aspect of [v]'s phonology, too. These

---

[23] Phonetic grounding of constraints in OT was a long-needed development in the
history of the theory. See Kager (1999, 11f), who warns against positing con-
straints "lacking both typological motivation and phonetic grounding, even if
there is compelling motivation for [them] from the language data under analy-
sis." Indeed, phonetically-grounded phonology can be seen as a reaction against
OT's highly suspicious arbitrary "parochial" and language-specific constraints.
See also Wright (2004) on this issue.

[24] This is why Petrova and Szentgyörgyi's criticism of Padgett's (2002) relegating
some phenomena of voicing assimilation in Russian to the phonetic component is
not just.

analyses included [v]'s distribution in (i) intervocalic CC clusters and (ii) word-initial obstruent + [v] clusters. In the first case, the fact that no [v]–voiceless obstruent clusters exist in Hungarian is due to the Voicing Requirement as well as that [v] can only stand before voiced obstruents. Word-initial obstruent–[v] clusters have also been widely discussed, due to their seemingly exceptional behaviour regarding Voicing Requirement. But no analysis we have reviewed has stepped beyond these cases and attempted to provide a **unitary** account of the distributional facts. Moreover, none of the past approaches have offered an explanatory analysis on the **frequency** facts concerning [v]'s distribution, namely that some clusters are very frequent, others are rare or even downright impossible (this we may call **phonotactic graduality**). For instance, the models focusing on voicing assimilation did not account for the fact that although [v]–voiced obstruent clusters are the only expected clusters (and not [v]–voiceless clusters), the number of such sequences is nevertheless extremely low.

Two well-established works whose pronounced aim is to present a rather comprehensive account of the phonotactic facts of Hungarian are Törkenczy (1994) and Siptár–Törkenczy (2000). Unfortunately, they cannot escape the problems we have mentioned in the previous paragraph: [v]'s distributional facts as well as its behaviour concerning voicing assimilation are not treated in a unified manner; in fact, both works need to resort to exceptional devices.[25] Furthermore, their account on [v]'s phonotactics is only descriptively adequate, but does not offer an explanation; the nagging problem of the inability to account for the graduality of [v]'s well-formedness in consonant sequences still remains, namely that the type frequency of the relevant clusters is varied and cannot be captured by categorical means.

4. A crosslinguistic overview: [v] in Czech, Slovak and Russian

The double-faced behaviour of [v] is not unique to Hungarian. On the contrary, it is striking in how many languages (which may not be related genetically) [v] displays a similar, recurring phonological patterning. Here we provide a brief overview of three (related) other languages, where [v]'s

---

[25] They include simple (intuition-based) statements like "we regard syllable-initial [vl] [...] as exceptional" (Törkenczy *op.cit.*, 282f), as well as exceptional representational measures such as the introduction of "Sp(ecial) Licensing" (see Siptár– Törkenczy *op.cit.*, 140).

behaviour has been reported to be very similar to that in Hungarian: these languages are Czech, Slovak and Russian.[26]

(24) The Voicing Requirement in Czech, Slovak and Russian

    (a)   obstruent + obstruent clusters: regressive voicing assimilation

        (i)   Czech: [s][d] → [zd]: *s domem* 'with a house'

                    [z][p] → [sp]: *z pole* 'from a field'

        (ii) Slovak: [s][b] → [zb]: *prosba* 'request'

                    [z][k] → [sk]: *z kina* 'from a cinema'

        (iii) Russian: [t][b] → [db]: *ot-brosit^j* 'to throw aside'

                    [d][p] → [tp]: *pod-pisat^j* 'to sign'

    (b)   obstruent + sonorant clusters

        (i)   Czech: sonorants do not trigger voicing assimilation

            [s][l] → [sl] (*[zl]): *s lesem* 'with a forest'

            [s][m] → [sm] (*[zm]): *s mužem* 'with a man'

        (ii) Slovak: sonorants can trigger voicing assimilation **across word boundaries**

            [p][m] → [bm]: *chlap môže* 'a guy can'

            cf. [p][m] → [pm] (*[bm]): *chlap-mi* 'guy.inst.pl'

            cf. [k][n] → [kn] (*[gn]): *vlákno* 'fibre'

        (iii) Russian: sonorants do not trigger voicing assimilation

            [k][n^j] → [kn^j] (*[gn^j]): *kn^{j}az* 'prince'

            [s^j][m] → [s^{j}m] (*[z^{j}m]): *pis^{j}ma* 'letter.pl'

    (c)   sonorant + obstruent clusters: no voicing assimilation in any of the three languages

(25) Word-final consonants and voicing in Czech, Slovak and Russian

    (a)   word-final obstruents are devoiced

        (i)   Czech: *mužem* [ʒ] 'man.inst.sg' ∼ *muž* [ʃ] 'man.nom.sg'

        (ii) Slovak: *zväzu* [z] 'union.gen.sg' ∼ *zväz* [s] 'union.nom.sg'

        (iii) Russian: *kniga* [g] 'book.nom.sg' ∼ *knik* [k] 'book.gen.pl'

    (b)   word-final sonorants are not devoiced in any of the three languages[27]

---

[26] On the Czech and Slovak data, see Hall (2003; 2004). The Russian examples were taken from Padgett (2002). Note that in his examples, he suppresses predictable palatalization, vowel reduction, and other irrelevant surface variation. For more on the Russian data and the exact domain of the voicing phenomena, also consider Halle (1959); Hayes (1984); Petrova (2003); Petrova–Szentgyörgyi (2004). Languages other than the ones we focus on have also been said to show similar [v]-patterns, they include Hebrew (Barkaï–Horvath 1978), Ukrainian (Cyran–Nilsson 1998), Swedish, Romanian (Lombardi 1995), and some dialects of English, too (Petrova–Szentgyörgyi 2004).

[27] For Russian this claim is not uncontroversial. According to Padgett (2002), however, the process is at best only gradient and optional, just as much as the infa-

(26) The Voicing Requirement and [v] in Czech, Slovak and Russian

    (a) [v] as potential target: regressive voicing assimilation (∼ obstruents)

        (i)   Czech: [v][p] → [fp]: *v pole* 'in a field'

        (ii) Slovak: [v][t] → [ft]: *v tom* 'in that'

        (iii) Russian: [v][k] → [fk]: *korovka* 'cow.dim'
            (cf. *korovok* [v] 'cow.dim.gen.pl')

    (b) [v] as potential trigger

        (i)   Czech: [v] does not trigger voicing assimilation (∼ sonorants)
            [t][v] → [tv] (*[dv]): *tvůj* 'your'[28]

        (ii) Slovak: [v] can trigger voicing assimilation **across word boundaries**
            (∼ sonorants)
            [t][v] → [dv]: *brat vám (zavolá)* 'brother (calls you)'
            cf. [t][v] → [tv] (*[dv]): *tvár* 'face'

        (iii) Russian: [v] does not trigger voicing assimilation (∼ sonorants)
            [t][v] → [tv] (*[dv]): *tver^j* 'Tver'

(27) Word-final [v] in Czech, Slovak and Russian

    (a) Czech: word-final [v] gets devoiced (∼ obstruents):
        zpěvem [v] 'song.inst.sg' ∼ *zpěv* [f] 'song.nom.sg'

    (b) Slovak: word-final [v] is realized as [w]:
        ústavom [v] 'institute.inst.pl' ∼ *ústav* [w] 'institute.nom.sg'

    (c) Russian: word-final [v] gets devoiced (∼ obstruents):
        prava [v] 'right.fem' ∼ *prav* [f] 'right.masc'

Clearly, [v] in Czech, Slovak and Russian displays the same behaviour as Hungarian [v] in consonant clusters with respect to voicing. There are three important differences we must point out, nonetheless. First, unlike these three languages, (Educated Colloquial) Hungarian has not been reported to show word-final obstruent devoicing (for [v] either). Second, Slovak appears to behave differently with respect to sonorant voicing: in this language, sonorants are said to voice voiceless obstruents before them across words and some morpheme boundaries. Lastly, Slovak is also special as in this language, word-final/pre-consonantal [v] is not realized as [f], but as [w], a type of lenition. The facts regarding voicing assimilation, word-final devoicing and the behaviour of [v] with respect to these two

---

mous transparency of Russian sonorants, as in *i*[s mts]*enska* 'out of Mcenk' vs. *i*[z o]*kna* 'out of the window'.

[28] In some dialects of Czech, [v] surfaces as a [f] after voiceless obstruents; this is the same behaviour as in the Western Transdanubian dialects of Hungarian (see (10)); further details can be found in Hall (2003; 2004).

phenomena are summarized in Table 4 for five languages (we also included
the Western Transdanubian dialect of Hungarian (WTH) in the chart;
"ECH" = Educated Colloquial Hungarian; "C" = Czech; "R" = Russian;
"S" = Slovak). A "+" is used if the relevant phenomenon applies in the
given language; "−" if it does not (but also consider the notes).

Table 4

Voicing effects and the behaviour of [v]
in Hungarian, Czech, Slovak and Russian

	ECH	WTH	C	S	R
regr. voice ass. in obs.+obs. clusters	+	+	+	+	+
regr. voice ass. in obs.+son. clusters	−	(+)*	−	(+)[†]	−
final obs. devoicing	−	+	+	+	+
[v] as target (devoicing)	+	+	+	+	+
[v] as trigger (voicing)	−	(+)*	−	(+)[†]	−
[v] in final devoicing	−[‡]	+	+	(−)[§]	+

* Sources do not specify the exact domain of presonorant voicing.
[†] Only applies across specific morpheme boundaries.
[‡] However, see the results of the experiment below.
[§] [v] is not devoiced but lenited to [w].

The most important generalizations are thus as follows: (i) all languages
display regressive voicing assimilation in obstruent–obstruent clusters;
(ii) target [v] behaves in all languages as an obstruent; (iii) [v] only trig-
gers voicing assimilation if the other sonorants do, too (and in the same
morphological context); (iv) word-final [v] patterns with obstruents with
respect to devoicing (with the exception of Slovak).[29] A last implication
that we can draw from Table 4 is that if a language displays presonorant
voicing (over word boundaries), then that language also has word-final
devoicing (the reverse is not true, however). The connection between
presonorant voicing and word-final devoicing seems to be an important
factor in languages where sonorants do not contrast in voicing but they
can be argued to demonstrate presonorant voicing.

---

[29] A more suitable generalization could be then to say that word-final [v] patterns
with obstruents in that it is "unstable": it is prone to changing in this position.
Actually, the analysis that we will be putting forth will predict what routes this
change can take.

In the remainder of the paper, we will try to argue for an analysis that can account for these recurring generalizations regarding [v]. Even though we will focus on Hungarian, we will predict that the account can be extended to the languages tackled in this section, and other languages, too.[30]

5. A phonetically-based approach to [v]

5.1. Functionalism and phonetics in phonology

There is a growing body of evidence that **functional** factors previously thought to be external to grammar can nevertheless exert direct influence on it. These factors include such "low level effects" as speech **production** (articulation) and speech **perception**.[31] The basic idea that we pursue in this paper is that the phonetic (specifically, the aerodynamic and acoustic) properties of sounds can regulate their phonological patterning (including segmental distribution and allophony). The four most important functional principles that we make use of here are listed in (28):

(28) (a) creation of contrast;

  (b) maximization of the number of contrasts;

  (c) maximization of the perceptual distinctiveness of contrasts;

  (d) minimization of articulatory effort.

The first of these principles is responsible for the creation of contrastive cognitive categories; by maximizing the number of contrastive categories (28b), the expressiveness of communication is enhanced by building up a substantial lexicon of categories. Principle (28c) accounts for the salience of the distinct basic categorical elements—according to it, categories must have acoustic properties that make them maximally different from each

---

[30] The works on [v] in the Slavic languages we have reviewed do not consider the **phonotactics** of [v], they only concentrate on its voicing behaviour (cf., however, Padgett's brief remarks on Russian obstruent–sonorant onset clusters (2002, 16)). We will, nevertheless, provide an account for Hungarian in which the phonotactic peculiarities can be explained as well. We anticipate that the analysis can also be transferred to the phonotactics of languages other than Hungarian.

[31] For a summary of functionalism, the role of articulation and speech perception in phonological theory, see, among others, Hume–Johnson (2001) and the references therein.

other perceptually; this principle thus requires segmental contrast to be sufficiently distinct and aims at **contrast preservation**. The last principle secures that the actual implementation (articulation) of the categories is to be carried out using as little energy as possible, a phonological consequence of which may be **contrast neutralization**. Notice that, as Flemming (2004) also shows, principle (28c) is inherently in conflict with both principles (b) and (d); thus, the phonological system is required to be such that it "weights" these principles and gives preference to one over the other.

The salience of a segmental contrast is **cue-based** and fundamentally depends on two factors: (i) the quality and quantity of the **inherent** acoustic cues of the given sound and (ii) the quality and quantity of the acoustic cues the sounds in its **immediate context** provide. Thus, the model makes no reference to segmental constituency (like the syllable), only the linear environment of a sound is considered. Segmental contrast is argued here to be licensed (upheld) in positions with sufficient and good-quality cues for that contrast to be perceived, recognized. Conversely, segmental contrast is predicted to be suspended, neutralized in badly-cued positions. This idea is expressed in the principle of **Licensing by Cue**:

(29) Licensing by Cue (Steriade 1999, 4):
    The likelihood that distinctive values of the feature F will occur in a given context is a function of the relative perceptibility of the F-contrast in that context.

Another phonetic notion that we introduce is the **articulatory target**, which is defined as a given configuration that the articulatory system has to reach, realize. In view of Licensing by Cue, we argue that a given target in a badly-cued context is not realizable, and depending on the specific phonetic characteristics of the target and its context, particular **articulatory adjustments** are made, with the consequence of contrast neutralization. Neutralization in a position is to be understood in two ways: either only one set of phonetic properties is realized there (say, voicelessness and friction), which the cues can license, or the segment is lost altogether (deletion).

Steriade (1997) argues that for a given contrast (such as voicing–voicelessness of obstruents), the environments the contrast may occur in can be scaled hierarchically according to how well or how badly they cue the contrast in question. For the voicing contrast of obstruents, she sets up the following hierarchy of environments (" $X < Y$ " means that Y is

an environment with less and worse perceptual cues than environment X (in other words, Y is more marked perceptually, less cue-able than X); "O" signals obstruents, "R" sonorants, and "#" a pause):

(30) Perceptual difficulty scale for the voicing contrast of obstruents:

   V__R < {#__R, O__R} < V__# < V__O < {O__O, O__#, #__O}

According to this hierarchy, obstruent voicing is best perceived between a vowel and a sonorant. It is in this position that the relevant cues for the contrast are available in number and quality: closure voicing, closure duration, the duration of the vowel, F1 values in the vowel, burst duration and amplitude, VOT value, F0 and F1 values at the onset of voicing in the second sonorant. However, the cues to obstruent voicing are curtailed after a vowel and before an obstruent/pause, where important transition cues on the right are now missing. In the worst contexts, in the lack of a neighbouring sonorant, (O__O , O__#, #__O), the all-important transition cues are unavailable, merely the inherent cues are kept (the voicing and duration of closure).

   We claim, along with Steriade (*ibid.*), that the hierarchy in (30) is "**inclusive**"; that is, if a contrast exists in a given position, then it also exists in those providing **better** cues for it. For instance, if a language has voicing contrast for obstruents in word-final position, it must also have that contrast after an obstruent and before a sonorant, word-initially, as well as between sonorants.[32]

5.2. The phonetics of voiced fricatives

We begin our phonetically-grounded, cue/target-based analysis of Hungarian [v] with considering the phonetics of voiced fricatives, in particular, their aerodynamic properties. For the articulatory system to target voicing and friction (turbulent noise) at the same time, an uneasy balance needs to be maintained.[33] First, high-amplitude turbulent noise requires a relatively high volume velocity of the air molecules as they pass through a channel. The faster the air moves, the louder the sound; thus, the ideal

---

[32] For more on the role of this monotonicity requirement in Hungarian and English phonotactics, see Rebrus–Trón (2002) and Kiss (2005).

[33] For the phonetic details of voicing and fricatives, see Stevens (1998, 477ff); Johnson (2003, 120ff); Jansen (2004, 40f) and Fuchs–Brunner (2005).

glottal configuration for a noisy fricative is when the glottis is **widely abducted** so that a high volume velocity airflow can pass through the vocal tract. Second, for a given rate of airflow, the narrower the channel, the louder the turbulent noise; according to Stevens (1998, 379ff), the glottal opening needs to be somewhat greater than the cross-sectional area of the supraglottal constriction. Lastly, turbulence is also produced when a jet of air hits an obstacle downstream. In the case of [s] and [ʃ], the upper and lower teeth constitute the obstacle, whereas for [f], it is the upper lip. If all these three factors are given, the sound produced is a high-intensity fricative.

On the other hand, for vocal fold vibration to take place, the glottis should be **closed** but loosely compressed, which is achieved by the air pressure to be **lower above the glottis** than below it.[34]

Based on the above, we can conclude that the production of high amplitude fricative noise and voicing involves contradictory articulatory gestures and aerodynamic events. Turbulent noise can be achieved by a widely abducted glottis and a constriction which is to be narrower than the glottal opening in the oral cavity. Both these factors result in an increase in intraoral pressure and a drop in subglottal pressure as more air builds up in the oral cavity at and behind the constriction area. An abducted glottis and a decrease in the transglottal pressure differential both remove the basic conditions for vocal fold vibration:

> "for the sake of continued voicing the oral pressure should be low, but for the sake of frication the oral pressure should be high. Meeting both of these requirements may be difficult. To the extent that the segment retains voicing it may be less of a fricative, and if it is a good fricative it runs the risk of

---

[34] As Jansen (2004, 36) estimates, the minimum transglottal pressure difference sufficient to maintain vocal cord vibration is around 200 Pa; to **initiate** fold vibration, this difference needs to be twice as much. A sound is defined to be **passively (modally) voiced** when during its articulation, a closed equilibrium position of the vocal folds and normal subglottal pressure is enough to initiate or maintain the physical conditions for vocal cord vibration (like in the case of sonorants). If passive voicing cannot be achieved (such as during the closure phase of stops), sounds are said to be **passively devoiced**. To overcome passive devoicing, a number of articulatory gestures (e.g., the enlarging of the oral tract volume) need to be implemented — this mechanism is referred to as **active voicing**, which characterizes ("truly") voiced obstruents. Conversely, **active devoicing** refers to situations in which sounds that would otherwise be passively voiced are nonetheless devoiced due to various additional articulatory means (e.g., the tensing of the vocal tract walls so that no passive expansion may occur and active decreasing of the size of the cavity behind the oral constriction by raising the larynx).

being devoiced. In fact, the noise component of voiced fricatives is much less than that for voiceless fricatives [. . .] and on nonsibilant voiced fricatives ([β, v, ð, j, ɣ, ʁ]) is often so weak as to be barely detectable."

(Ohala 1983, 201f)

In our analysis of Hungarian [v], we take this line of reasoning a step further. Our starting point is the aerodynamic incompatibility of the articulatory targets: our hypothesis is that the **maximal** implementation of **both** active voicing and strong frication at a labiodental place cannot be realized. The two targets can only be maintained if (i) voicing is passive/modal, due to a relatively open vocal tract, and (ii) frication is relatively low as a result of a wider constriction (much wider than that found in voiceless fricatives). The simultaneous realization of these targets can optimally be upheld in intervocalic, or more generally: intersonorant position, where passive voicing passes through the sequence of segments unimpeded (as in say [ava] or [avla]). Along these lines we claim that [v] is like a sonorant with respect to its voicing qualities but like a fricative in possessing turbulent noise, too, albeit at a lower level. This is in line with Padgett's (2002) analysis of Russian *v*, which he calls a **narrow approximant** and transcribes it as [ʋ̞] (from now on, we also adopt this symbol for notating Hungarian (prevocalic) *v*).[35] With respect to its noise qualities, this sound is claimed to stand between actively voiced and noisy [v] and passively voiced approximant [ʋ/w]. Between two sonorants, [ʋ̞] is predicted to display more formant structure and intensity and less turbulent noise than any other fricative (thus its noise is modulated by voicing), but less formant structure and intensity and more turbulent noise than any other ("wide") approximant.

Even though we base the starting point of our present analysis of [v] on aerodynamic grounds, a plausible perceptually-based account can also be suggested, as is done in Balise–Diehl (1994). They claim that the presence of voicing interferes with the perception of place cues in fricatives. Voicing-based laryngeal contrasts in fricative inventories tend to be neutralized because it is relatively hard to recover their place cues. They cite two pieces of evidence in support of this: the presence of voicing in a fricative reduces the amplitude of frication noise, which is an important cue for place contrast; furthermore, studies of consonant confusions indicate that across various signal-to-noise ratios, voiceless fricatives are identified correctly more often than their voiced counterparts.

---

[35] See also Lulich (2004), who arrives at a similar conclusion concerning Russian [v], also working in a phonetics-based model.

In lack of additional and detailed data on place cues, we, nevertheless, hypothesize that [ʋ̠] in certain positions is (partially) neutralized because the relative perceptibility of its target cues—**voicing** and **noise**—, which are to be perceived simultaneously, is severely curtailed. In accordance with the functionalist stance we briefly introduced in the previous section as well as the aerodynamic premise presented above, we predict that [ʋ̠] in an unfavourably cued context can take two partial neutralization routes. Since noise (turbulence) and voicing cannot simultaneously be maintained in such environments, either (i) noise is preserved and voicing is lost, or (ii) voicing is kept and noise is "sacrificed". In the first scenario— when there is no vocal fold vibration and hence the glottis is open—the aerodynamic premise predicts that a rather noisy devoiced fricative ([v̥]) is produced as more air pushes out at a given (and relatively constant) constriction degree, the result of which is considerably more frication. [ʋ̠]'s fricativization under devoicing is in line with the behaviour of other approximants; for instance, [l] and [j] (with wide "constriction") also show frication (thus occur as [ɬ] and [ç]) when they occur in positions where they devoice.[36] In our approach, a necessarily devoiced narrow [ʋ̠] should thus show more friction than devoiced wide approximants. In the second case, when constriction is necessarily less narrow, friction is lost but voicing is maintained, and the sound produced is a labial (wide) approximant, a [ʋ] or [w].

In accordance with Steriade's (1997) hierarchy of environments based on direct reference to the number/quality of acoustic cues to the voicing contrast of obstruents (see (30)), we also claim that the same hierarchy is applicable to the explanation of the neutralization effects of [ʋ̠]. Thus, [ʋ̠]'s cues to voicing (and friction, place, etc.) are best perceptible between two sonorants. The moment [ʋ̠] finds itself before a non-sonorant (an obstruent or silence), its voicing cues are severely reduced in number and quality, and so it is in this environment that the two neutralization routes are predicted (in case the segment is preserved). We assume that languages are free to choose which neutralization strategy they follow: devoicing/strong friction or de-noising/vocalization. The first strategy is selected by languages like Hungarian, Russian; whereas the second one is chosen by Slovak, for instance. Absolute loss of contrast, namely segment deletion, is also a possible option in this model; actually, the

[36] See Padgett (2002, 22) for examples from Norwegian, Iberian Spanish and French. Hungarian [j] also displays a similar conduct in neutralization-prone contexts, such as after an obstruent and before a pause: *lépj* [leːpç] 'step.imp'.

v-deletion strategy was apparent in some of the cases of the experiment we conducted (see section **5.3.2**).

Our predictions as to the (likely) **realization** of Hungarian [ʋ] in the various contexts are shown in (31) (we use the same hierarchy of contexts as in (30)):

(31) context: V__R │ #__R, O__R │ V__# │ V__O │ O__O, O__#, #__O
      realiz.: [ʋ̝] │   [ʋ̝]   │ [ʋ̥] │ [v̥] │   [v̥]

Thus, Hungarian [ʋ] is licensed only before sonorants (including vowels), in all other contexts, it is likely to devoice and obstruentize. We must note that these predictions on [ʋ]'s realizations are founded on the phonetic (aerodynamic/cue-based) premise alone; as we will see, other functional factors (such as the possible coarticulatory effect of the active voicing of a following obstruent) can modify these predictions.

Our cue-based approach can also predict the **phonotactic patterning** of [ʋ]. Since its contrast receives the most salient cues before a sonorant, it is these types that we expect to display the most items (thus words with pre-sonorant [ʋ̝]). As we move down on the scale of environments towards those with less and less cues for the recognition of [ʋ]'s contrast, we anticipate fewer items to contain this consonant; actually, this model predicts that it is when [ʋ̝]'s phonetic targets cannot be realized (i.e., when it is articulated as [v̥]) that its distribution is restricted. Thus, this analysis predicts what surface variants can appear in the various contexts if a specific set of phonetic targets is to be realized, as well as [ʋ]'s gradual phonotactic patterning: both aspects are intimately linked and directly grounded in phonetic factors.

Furthermore, [ʋ̝]'s behaviour with respect to the "**Voicing Requirement**" can also be explained on the phonetic grounds that have been tackled here. In pre-sonorant position, as we argued, [ʋ] is realized as a passively voiced (narrow) approximant (in other words, it lacks active/ dynamic voicing targets). As such, it is expected to pattern with other sounds that bear similar voicing characteristics, like sonorants. It is for this reason that pre-sonorant [ʋ̝] will not voice a previous obstruent. Obviously, the question fundamentally boils down to the phonetics of active/ passive (de)voicing and voicing assimilation. Lack of space prevents us from providing a detailed phonetics-based analysis on why sonorants do not normally actively voice obstruents, and why they appear to nevertheless do so in some languages; suffice it to say that we see the core of the problem to be related to the **coarticulatory** properties of voicing targets

(in this we follow the ideas put forth in Farnetani 1997 and Jansen 2004, among others). Only actively voiced and devoiced sounds are assumed to participate in voicing assimilation as their voicing/devoicing-enhancing gestures can "spill over" into neighbouring segments (mainly those preceding them). Passively voiced sounds, on the other hand, do not possess voice-enhancing gestures, and so they "can have no coarticulatory effect on the voicing control of neighbouring obstruents: [...] there is simply nothing to spill over into flanking sounds" (Jansen *op.cit.*, 108).[37]

When [ʋ] is not followed by a sonorant but by a **voiceless obstruent** (as in *savtól* 'acid.abl'), the coarticulation-based voicing assimilation model (together with the cue-driven approach we propose) predicts that [ʋ̥] should appear as a truly voiceless noisy fricative. This is because voiceless obstruents are claimed to be actively devoiced in Hungarian with devoicing gestures that can spill over into [ʋ]. Remember that our model predicts that [ʋ] should be realized as a devoiced and noisy sound ([v̥]) before obstruents; we assume that the active devoicing gestures of a following voiceless obstruent only enhance the voicelessness of [v̥], and so we expect a sound very close to [f] to be realized, with the consequence of the [f]–[ʋ̥] contrast to be (completely) neutralized. We assume thus that *v* is extremely sensitive to the nature (obstruency, voicing, etc.) of the segment following it (and possibly of that preceding it) because its targeted articulatory gestures are difficult to maintain at the same time, and can only be upheld under optimal circumstances (next to sonorants). Consequently, the model hypothesizes that if *v* is followed by a **voiced consonant**, it will more easily receive voicing from it, and depending on the aperture qualities of that consonant, *v*'s realization will gradually move between stages of (i) a weakly fricated voiced narrow approximant [ʋ̝] (before vowels and sonorant consonants), (ii) a more fricatival voiced

---

[37] It follows from this line of reasoning that only those languages are expected to display voice assimilation to sonorants where sonorants contrast in active voicing and active devoicing. It is likely that languages that do not contrast voicing in sonorants (thus they are passively/modally voiced) can only show voicing assimilation effects if sonorants target passively (de)voiced/neutralized obstruents, which thus lack targets for voicing (as in the case of word-final neutralized obstruents in Slovak and the Western Transdanubian dialect of Hungarian; see above). As Jansen argues, "if neutralized obstruents indeed lack voicing targets, they should show a greater degree of voicing between a vowel and a following sonorant than actively devoiced obstruents, simply as a result of the **passive continuation of voicing into the constriction phase**. It could well be this increased amount of voicing (relative to utterance-final and [voiceless] contexts) that is interpreted by linguists as voicing assimilation" (*op.cit.*, 119; our emphasis). Further quantitative phonetic research is obviously required here.

narrow approximant/fricative, thus a sound close to [v] (before voiced
obstruents or even narrowly produced approximants) to (iii) a very noisy
devoiced fricative [v̥/f]. Accordingly, the modified table on the predictions
concerning v's probable realizations in Hungarian is (32):[38]

(32) context: V__R │ #__R, O__R │ V__# │ V__O │O__O, #__O, O__#
        realiz.:   [v̞]   │ [v̞]   [v̞] or [v̥] │ [v̥] │[v̥/f] or [v]│   [v̥/f] or [v]   [v̥]

In the following section, we will present the results of an acoustic ex-
periment that aimed at checking the various phonetic properties of the
realizations of Hungarian v, and hence the validity of the predictions of
the phonetically-based model we have introduced here.

5.3. Hungarian [v] realizations—an acoustic experiment

5.3.1. Method

In the experiment to be presented here, we focused on the acoustic real-
ization of v next to a consonant. We examined vC and Cv clusters word-
initially, word-finally and in intervocalic position; the study included VvV
sequences as well, as a point of reference, since we consider the realiza-
tion of v in this position as the prototypical manifestation of what we
described as the narrow labiodental approximant [v̞]. Where it was pos-
sible, we only investigated monomorphemic forms (hence to cover the
items of Table 1); in some cases, however, we looked at words where a
morpheme boundary divided the members of the clusters in question (as
in *hívj, hívd* for instance).

Ten native speakers (six female and four male) of Educated Col-
loquial Hungarian were asked to read out the test sentences (see the
Appendix) at a normal speech rate in a sound proof cabin. The age of
the speakers were between 22 and 28, with the exception of two sub-
jects, who were aged 58 and 60 respectively. They were not paid for the
participation in the experiment. The data were recorded with a Sony
ECM-MS907 microphone onto a Sony MDMZ0710 minidisk, digitized at
44100 Hz and resampled at 22050 Hz. The acoustic measurements were
computer-analysed using Praat (Boersma–Weenink 2005).

The experiment aimed to measure the following parameters:

[38] We would like to stress again that the scale predicts a **probability** of v-realizations
(given a normal speech rate/style). The "boundaries" of where the actual real-
izations may occur are not strictly predicted, and may shift.

(33) (a) spectral moments, primarily **Centre of Gravity (CoG)** (but also skewness and standard deviation)

   (b) **voicing** (i.e., glottal pulses)

   (c) **formant structure** (transitions from neighbouring segments as well as in the steady state of the segment)

   (d) segmental **duration**

Here, we will mostly focus on CoG and voicing parameters.

Spectral moments have been widely used in the literature (Jassem 1979; Forrest et al. 1988; Ladefoged 2003, 156ff; Gordon et al. 2005; Machač–Skarnitzl 2005) to quantify consonantal — in particular fricative — characteristics, especially with the need to distinguish one fricative from the other, across vowel context and speaker.[39] The first spectral moment, CoG (or "centroid"), is a measure for how high the frequencies in a spectrum are on average, it thus represents a spectral mean. CoG is one of the reliable indicators of turbulent noise: the higher the CoG value, the noisier the segment. For instance, for a white noise sampled at 22050 Hz the CoG value is 5512.5 Hz, i.e., half of the Nyquist frequency. It is important to note that there are important language-specific and individual variations in the CoG values of a given fricative; the relative order of the CoG values of different sounds is what is of interest in this paper. The value of (normalized) skewness also indicates how noisy a sound is, it shows how much the shape of the spectrum below the CoG is different from above it, for a white noise skewness is zero. Standard deviation is to be interpreted as a measure for how much the frequencies in a spectrum can deviate from the CoG (for a sine wave, the standard deviation is zero, for a white noise, it is the Nyquist frequency divided by $\sqrt{12}$). In those cases when v is realized less fricatival and more approximant-like, the CoG values are less informative, vowel transitions, therefore, the formant structure of the segment are of more interest.[40]

Voicing was simply measured on the basis of periodicity in the waveform as well as the presence/absence of a voice bar in the spectogram.

---

[39] See, however, Shadle–Mair (1996) and Shadle (2005) on the problems of using spectral moments in reliably quantifying the **differences** between fricatives. They can, however, be reliably used to quantify turbulence from non-turbulence, which is the prime motivation of this paper.

[40] As we have indicated, we primarily focus on the first spectral moment, the CoG, in this paper. Nonetheless, the other spectral moments we examined, skewness and standard deviation, were fully compatible with the CoG values we obtained. For instance, whenever CoG was relatively high, skewness approximated to 0 Hz.

Formant transitions were also checked, in particular for results with low CoG (which thus suggested the presence of a less turbulent segment). Relatively clear formant structure and transitions were indicative of a more vowel-like segment.

Measurement samples were gained the following way. The boundaries of v were determined by visual inspection of spectograms and waveforms and listening to the recordings. The selections containing the v portions contained "extra" parts of at least 25–25 ms at the edges, as these areas were not used by the analyzer program due to the analysis window shape and size. Thus, the usable area safely included the whole v portion. Male voices were band filtered at 0–300 Hz, female voices at 0–400 Hz. Several fast Fourier transform (FFT) spectra were computed for each v using a Gaussian window shape with a physical length of 50 ms (which is the sum of the "effective" lengths of 25–25 ms around the cursor in the analysis window). The first measurement was taken at 28 ms, then the 50 ms window was shifted with 5 ms increments until the whole of the segment was covered. In this way several overlapping slices were made of each v-segment, depending, of course, on the duration of the given segment sample: for shorter v-instances (below 50 ms), a 30 ms Gaussian window was employed (with 15–15 ms around the cursor) and only one or two slices were gained. The multiple windowing method is indispensable in the spectral analysis of sounds (such as fricatives) whose distribution of energy is (quasi-) random, inconstant.[41] Lastly, the CoG, skewness, standard deviation and formant values (F1 through F5) of each slice were computed and finally averaged. Thus, averages were obtained for each v occurrence for each speaker, but we also averaged every v occurrence for all speakers.

5.3.2. Results

Figure 1 exhibits the averages of the CoG values for all speakers. Axis y shows the CoG averages, whereas axis x contains all the words we tested in the experiment.

The results of the averages of the CoG values of all test words back up the main hypothesis: v occurring in a considerably unfavourable context—for instance word-finally preceded by a consonant—becomes very noisy/fricatival (see the results for *könyv, ölyv, jókedv, nyelv* etc. on

---

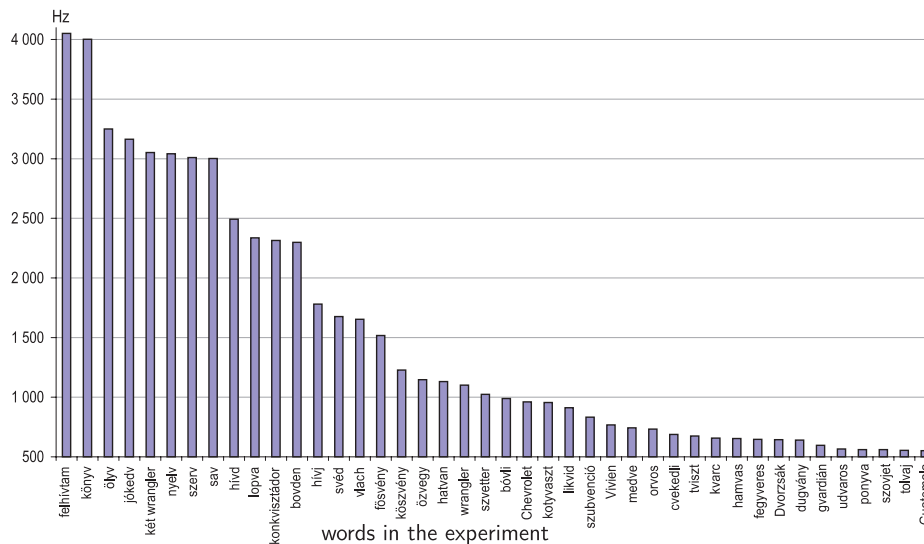[41] On the problems of the spectral analysis of turbulent sounds, see Ladefoged (2003, 153ff).

Fig. 1

CoG averages (all speakers)

the left of the figure: the average CoGs of these *v*'s are between 3000–4000 Hz). In a somewhat more favourable context, when *v* is preceded by a vowel but not followed by one, as in *sav* and *hívd*, we still received high CoG values: *v* was realized with a fair amount of fricative noise (2500–3000 Hz). Low CoGs occurred in pre-sonorant position. This is exactly what was indicated concerning *v*'s noisiness in earlier accounts as well (as in Siptár 1996; Siptár–Törkenczy 2000; see (3) on page 183).

It is noteworthy however, and has not yet been noted in the literature to the best of our knowledge, that in cases when *v* is realized as a strongly fricated sound, **it often loses voicing** at the same time. Thus high CoG goes hand in hand with devoicing. This is displayed in Figure 2 (p. 210).

Axis *y* of Figure 2 shows the number of subjects participating in the experiment, axis *x* contains the words tested in a decreasing order of their CoG values. The results back up our hypothesis according to which, *v* can preserve both its friction and voicing targets (as well as place) in pre-sonorant position; in other (unfavourable) positions, it rather tends to lose its voicing and become strongly fricated. Relatively high CoG and (simultaneous) voicing-preservation was observed in pre-obstruent position, where the obstruent is voiced (as in *bovden*). The highest level of CoG and devoicing were found in cases where *v* stands before a voiceless
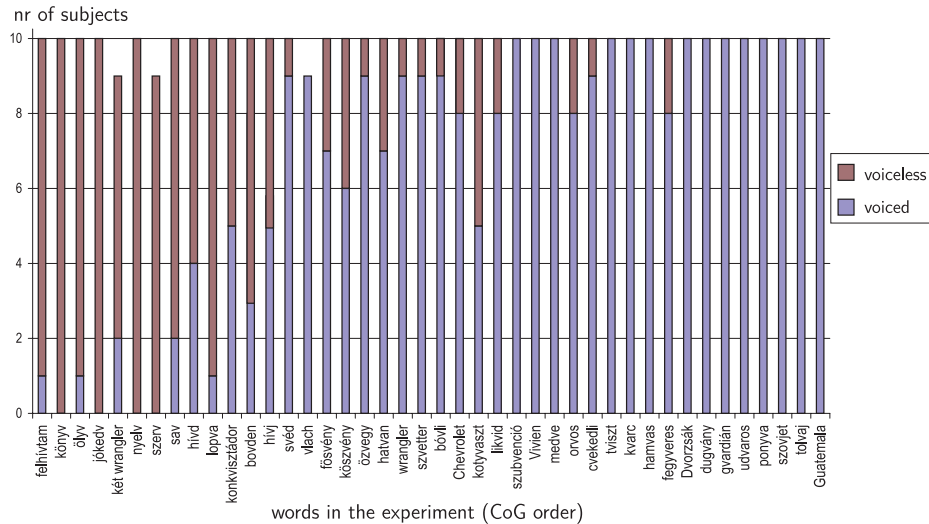
nr of subjects

Fig. 2

CoG and voicing

obstruent (as in *felhívtam*), actually, the CoG values of this particular realization is basically the same as that of [f]. Thus, in accordance with the predictions (see (31)), we found the following three-way *v*-realization pattern:

(34) (a)   strongly fricated and devoiced [v̥/f] (e.g., *felhívtam, könyv*)

   (b)   moderately fricated and voiced [v] (e.g., *bovden, bóvli*)

   (c)   weakly fricated and voiced [ʋ] (e.g., *udvaros, Vivien*)

Figure 3 provides spectograms (with corresponding waveforms) and FFT spectra, taken from the middle of the non-transitional portion of *v*, for each realization in (34). Linear predictive code (LPC) smoothed spectra are superimposed in the FFT spectra for easier viewing. We also added an example of a word-initial [f] (*falnak* 'wall.dat') for the sake of better comparison. The similarities between the acoustics of this [f] and the *v*'s of *felhívtam* 'I called' and *könyv* 'book' are striking (see (a–c) on page 211). None of these segments are voiced (see the lack of vertical stripes in the spectograms and the aperiodicity on the waveforms) or contain formant structure (cf. the abrupt start/finish of formant transitions at neighbouring sonorants), and all three show flat/peakless spectra typical of (diffuse)
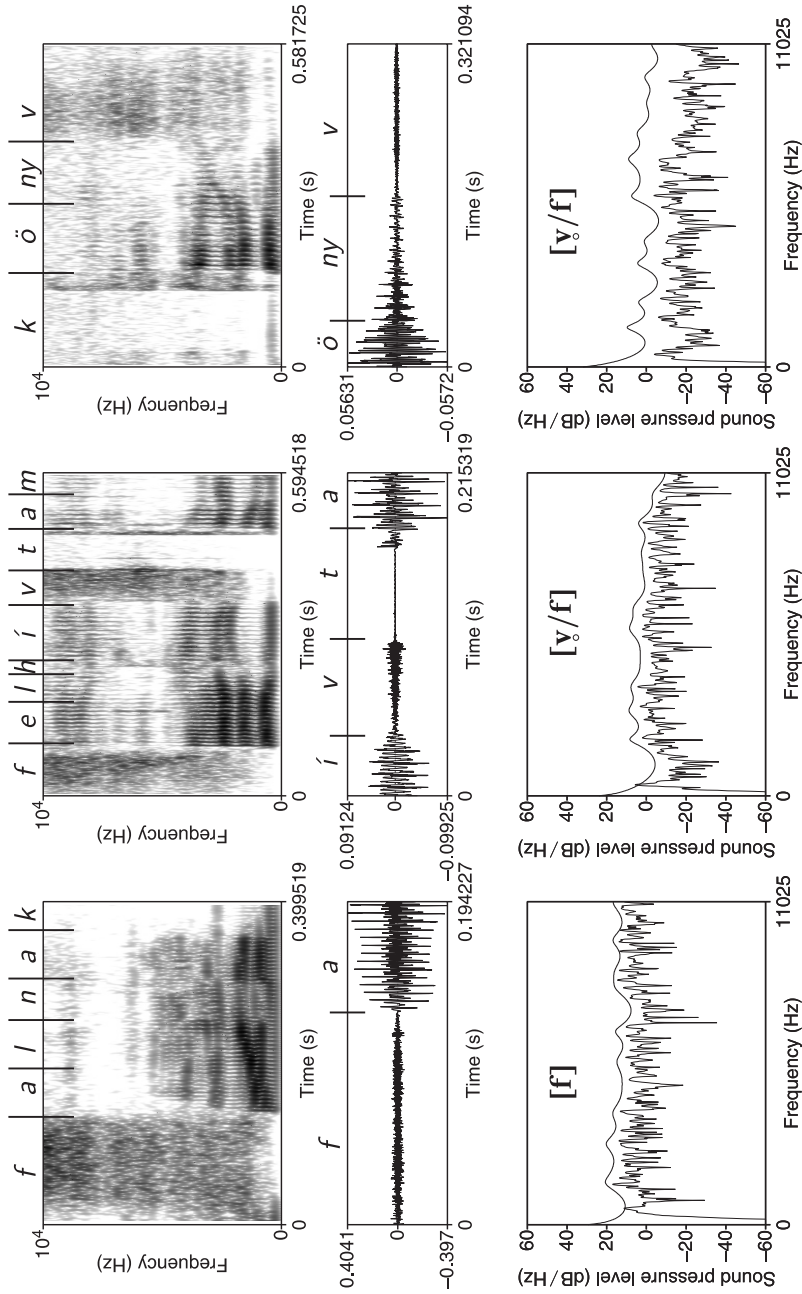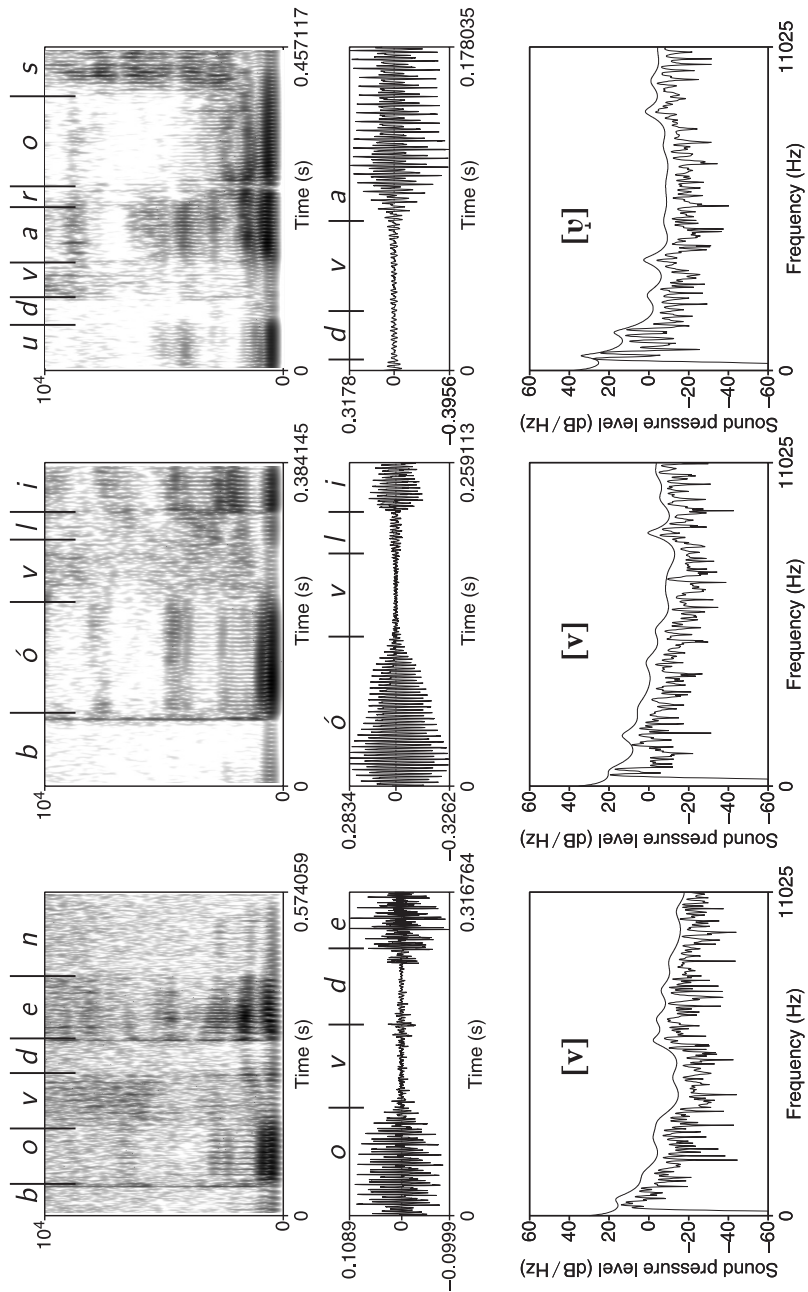
Fig. 3
f and v-realizations (1)

(a) falnak (subject 04)

(b) felhívtam (subject 01)

(c) könyv (subject 03)

(d) *boyden* (subject 10)

(e) *bővli* (subject 07)

(f) *udvaros* (subject 04)

Fig. 3
v-realizations (2)

labial/labiodental voiceless fricatives with energy spread over a large frequency range. We also found some correlation—although the durational aspect of v was not the main focus of our study—between length and friction. When v was realized as a noisy fricative, it was usually long (around 100 ms), and when it was realized as an approximant, it was usually less than 50 ms long, sometimes even less than 30 ms (this was usually the case in C__V, where the release of the C often masked v, and the formant transitions of the V passed through it, as it were).

Examples are given for the v's in *bovden* 'V-shaped belt', *bóvli* 'trash' and *udvaros* 'with a courtyard', respectively in (d–f) (on page 212). Both *bovden* and *bóvli* in these particular instances are voiced but they are also somewhat noisy (not as much, however, as the previous instances of v in (a–c)), which can also be seen in their spectra: they are not as diffuse, in fact, lower frequency regions show some intensity. It is these segments that we consider relatively fricated as well as voiced (hence a [v]). The last example (*udvaros*) shows that v is voiced (cf. the glottal pulses in the spectogram and the periodicity of the waveform), has a much clearer formant structure continuing into the vowel following it, and its spectrum exhibits characteristic vowel-like peaks at low frequencies. This segment is thus what we described as the narrow labiodental approximant [ʋ].

In the "harshest" context (C__#, especially O__#), we always obtained high CoGs and devoicing for v, as exactly predicted by the hierarchy of cues in (30), approximant [ʋ̥] was never produced here (one subject deleted v in *szerv* 'organ', too). However, this was not the case for the environment V__O: the averages in Figure 1 must be taken with a pinch of salt as they conceal some important generalizations. It is not true that all occurrences of v in, e.g., *bovden* were fricated (with an intermediate CoG of around 2000–2500 Hz) and voiced. In fact, and this is evident in Figure 2, seven subjects pronounced v here devoiced. Actually, five subjects even devoiced the following d, too. Crucially, however, **whenever v was devoiced, it was also fricated**, as evidenced by the relatively high CoG averages. The CoG–voicing data of pre-consonantal v (as well as of the consonant) for each subject are displayed in Table 5 (overleaf). The data in the table back up our earlier predictions, namely, the exact nature (obstruency/voicing) of the consonant after v does influence its realization. It looks as if v is most unstable before the voiced obstruent d and most stable before j ("stable" referring to the fact that v can preserve all its articulatory targets). l and r occupy an intermediate position in affecting v. Before d, v is more likely to occur as

Table 5

Average CoG and voicing of pre-consonantal *v* (1)

	bovden				*Chevrolet*				*bóvli*				*szovjet*		
s	CoG	vo-v	vo-C	s	CoG	vo-v	vo-C	s	CoG	vo-v	vo-C	s	CoG	vo-v	vo-C
1	2787	0	1	1	1555	0	0	1	2443	0	1	1	450	1	1
2	896	1	1	2	786	1	1	2	739	1	1	2	627	1	1
3	775	1	1	3	1399	0	1	3	534	1	1	3	497	1	1
4	1992	0	0	4	638	1	1	4	644	1	1	4	692	1	1
5	4619	0	0	5	769	1	1	5	1396	1	1	5	618	1	1
6	3621	0	0	6	2070	0	0	6	1074	1	1	6	708	1	1
7	2904	0	0	7	802	1	1	7	949	1	1	7	486	1	1
8	2372	0	0	8	551	1	1	8	951	1	1	8	594	1	1
9	2194	0	1	9	583	1	1	9	646	1	1	9	483	1	1
10	824	1	1	10	454	1	1	10	508	1	1	10	443	1	1

Legend: "s" = subject; "CoG" = average CoG value (Hz); "vo-v" /
"vo-C" = voicing of *v*/following consonant (0 = voiceless, 1 = voiced)

devoiced/strongly fricated ([v̥]). Only two subjects maintained voicing in
the cluster (but then, *v* lost its noisiness by a considerable degree). It is
definitely compelling that before *d*, *v* prefers to lose voicing to actually
receiving the voicing of the obstruent. We speculate (and consequently,
a more thorough research is required) that obstruency is weighted more
in the partial neutralization of *v* rather than the coarticulatory spill-over
of obstruent voicing: therefore, regardless of *d*'s voicing, this context still
counts as partially neutralizing for *v* to a [v̥]. Remember: the marked-
ness of the pre-*d*, and generally pre-obstruent, context is also manifest in
phonotactics: such sequences are extremely rare or non-existent. *j*, just
like vowels, does not make *v* unstable: here it preserves both its voicing
and weak frication, and so it occurs as a [ʋ]. With respect to the CoG and
voicing parameters, *r* is more like *d*, while *l* is more like *j*. Even though a
sonorant, *r* makes *v* lose its voicing (and get fricated) more often than *l*.
We assume that this may be due to the trilled articulation of Hungarian
r, namely, that its articulation involves complete (albeit short and rapid)
closures. Nevertheless, we conclude that in this respect, *r* can constitute
an unfavourable context for *v*. (Notice that two subjects even devoiced *r*
when *v* was devoiced.) Before *l*, the voicing of *v* is preserved; we, however,
received a little higher degree of frication than before *j* and vowels. Ac-
tually, it was here that we obtained the highest CoG values while voicing
was also preserved (see subjects 5 and 6).

Further data also back up our hypothesis that the immediate context of *v* has a great impact on its realization. When the consonant following it is itself in a context marked aerodynamically/perceptually, this impoverished consonant affects *v*, too. Two such cases were measured (*hívd* 'call.2sg.def.imp.', *hívj* 'call.2sg.indef.imp.'), the results are shown in Table 6.

Table 6
Average CoG and voicing of pre-consonantal *v* (2)

	hívd				*hívj*		
s	CoG	vo-v	vo-C	s	CoG	vo-v	vo-C
1	4203	0	0	1	4038	0	0
2	496	1	0	2	2492	0	0
3	1607	1	0	3	2630	0	0
4	1992	1	0	4	762	1	0
5	2950	0	0	5	759	1	0
6	4044	0	0	6	2546	0	1
7	3014	0	0	7	2474	0	0
8	2882	0	0	8	669	1	0
9	840	1	1	9	954	1	0
10	2903	0	0	10	487	1	0

Under such extreme circumstances (i.e., word-finally, after another consonant), *d* and *j* tend to become unstable: both were devoiced in 9/10 cases (*j* actually also became highly fricated when it was devoiced, just like *v*). When these segments were devoiced, *v* was devoiced/fricated in five out of nine cases. When *v* preserved its voicing, its CoG varied, indicating a wide-to-narrow constriction. But whenever we obtained high CoG around 1600–2000 Hz as well as voicing, it involved what we may describe as "careful" pronunciation.

To sum up, we can conclude that the environment C__# constitutes one of the worst contexts for maintaining all the articulatory targets of *v*: it is in this position that *v* is most likely to lose its voicing and become very strongly fricated (in Hungarian at least), **no other alternative is allowed** (disregarding careful speech styles, of course). The environment V__C, on the other hand, is one notch better and thus occupies an intermediate position in the context hierarchy: as we have seen, depending primarily on the characteristics of the following consonant, *v* ([ʋ]) can fully be maintained (with low CoG and voicing) **or** lose its voicing and

become fricatival; here thus, due to the intermediacy of the context, the two alternatives that our model predict are indeed possible.

Let us focus now on the **post-obstruent** (and prevocalic) position. This context also showed the same two-fold variation as the pre-consonantal context. *v* either preserved its voicing and was weakly fricated or it was devoiced and rather fricatival. High CoG almost always indicated devoicing, and visa versa. Again, however, the nature of the consonant does seem to play an important role: some consonants are more likely to cause devoicing of *v* than others. For example, in V*tv*V clusters, *v* kept its voicing and weak frication (7/10 subjects) more often than in V*pv*V clusters (3/10). Very similar results were obtained for V*kv*V clusters: in this context, *v* kept its voicing more often. Furthermore, it seems that the fact that the cluster is word-initial is also of importance: word-initially, 10/10 subjects pronounced a voiced *v* with a CoG less than 1000 Hz after *t*, and 9/10 after *k*. Figure 4 shows an example of the two strategies in the pronunciation of *hatvan* 'sixty' (voicing in the spectograms below is also indicated by dotted lines).
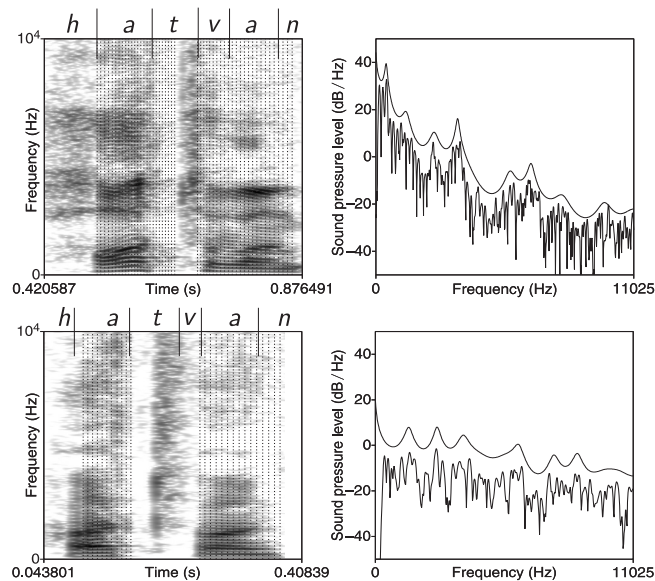

Fig. 4

Realizations of *v* in *hatvan* 'sixty'
(subject 3 vs. subject 10)

The first spectogram and spectra (of subject 3) show a typically approximant-like realization (consider the scarcity of noise at high frequencies, the clear formant structure throughout v and the following vowel (a), the voicing bars in the spectogram, as well as the formant peaks at low frequencies in the spectrum); the second spectogram and spectra (of subject 10) show that v is voiceless, noisy at a wide range of frequencies and lacks well-definable formant structure.

The realizational patterns can thus explain the apparent puzzle of why there is no voicing assimilation in voiceless obstruent–v clusters: when v is realized as a narrow approximant, its passive voicing cannot spill over to the preceding obstruent, when, however, it devoices and fricates (as in *lopva* 'stealthily'), it being voiceless, there is nothing to voice.

The influence of surrounding consonants on v is once again evident in **voiced obstruent–v clusters**: in this case, v almost never got devoiced, words like *Guatamala, udvaros, dugvány, Dvorzsák, medve, szubvenció* were all pronounced with low CoGs and voiced.

We only noticed devoicing when a **voiceless** obstruent (in some cases r, too) stood before v and it was pronounced with a relatively strong/ long release. The masking effect of the turbulence of release is, we hypothesize, the reason behind the devoicing/frication of v in some of the cases. Actually, such masking effects were also observable in [z]v, [s]v and [ʃ]v clusters: here, the fricative noise spilled over into v ([z] actually devoiced too in cases like these). We present here two (partial) spectograms showing the two manifestations of v in *fösvény* 'miser' as pronounced by subjects 1 and 3 in Figure 5.
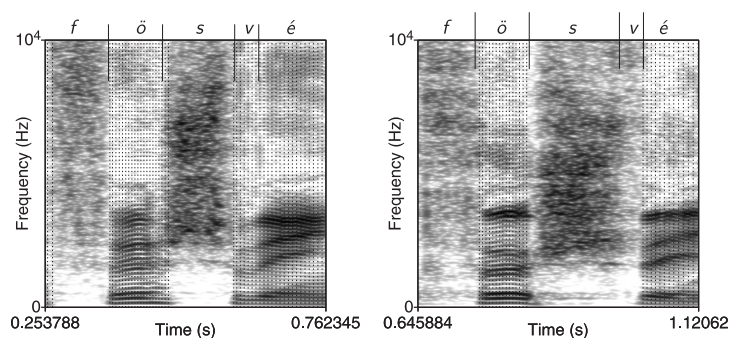

Fig. 5

Realizations of v in *fösvény* 'miser'
(subject 1 vs. subject 3)

In the first spectogram, there are unimpeded formant transitions in v continuing into the following vowel, v is also voiced; in the second, the noise of [ʃ] gradually extends into v, the vowel after it begins abruptly; v is not voiced here.

We can thus conclude that coarticulation effects play a significant role in the realization of v (and hence its phonological behaviour) in C–v clusters, too. Further research must find out the exact phonetic motivations. We see place of articulation and its relation to release noise/voicing to exert an important influence on v, whose target gestures we argue to be very instable (and fundamentally contradicting) and predisposed to coarticulatory effects. (Lulich's (2004) findings on Russian [v] also appear to support these ideas.)

The realizations of v in the remaining set of words, that of **word-initial #vC clusters**, also conform to our previous predictions. This context is, again, a highly infelicitous environment to preserve v's targeted noise/voice gestures. This is also shown by the rarity of such clusters in Hungarian (under this analysis an expected result). Their scarcity as well as the fact that they are foreign proper names made the testing rather difficult and therefore conclusive conclusions are hard to draw. It was nonetheless precisely in these tokens that we observed the most variability in the realization of v among the subjects. The following four "strategies" were noticeable: (i) some subjects attempted to pronounce these words rather slowly/carefully (as if putting them into "phonological quotes"); in these cases, both voicing and some noise were preserved; (ii) a few subjects pronounced these v's devoiced and with a turbulent noise; (iii) in the case of *Wrangler*, v was pronounced by some subjects as if it was an English [w] (even though, of course, no [w] occurs in this word in English), and lastly (iv) we also observed v-deletion (hence a complete loss of contrast) in the word *Vlach*. Actually, whenever v was preserved as voiced, it was extremely short in all cases (30–35 ms), which made segmentation very difficult.

Two example realizations of v in utterance-initial *Vlach* are displayed in Figure 6. In the first spectogram and spectra, we can see a very high degree of turbulence (incidentally, this token showed an average CoG of 4500–5000 Hz) and no voicing whatsoever; the second spectogram and spectra show a very vowel-like realization: notice the distinct and continuous formant structure, the lack of turbulence (average CoG: 559 Hz), and voicing in the spectogram, as well as the intensity peaks at low frequencies in the spectrum. Also take notice of the durational differences between the two v realizations.
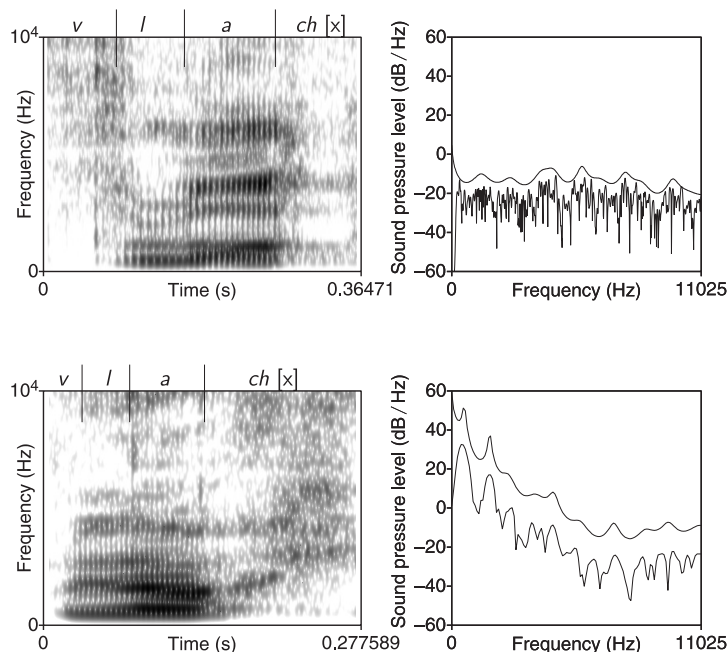
Fig. 6

Realizations of *v* in utterance-initial *Vlach*
'Vlachian/proper name' (subject 3 vs. subject 6)

We also tested *Wrangler* following voiceless *t* in the phrase *két Wrang-ler* 'two pairs of Wrangler jeans'. The salience of *v* in this position is rather severed; remember: this is actually one of the worst contexts in the cue hierarchy. And considering in particular the obstruent-like mani-festation of Hungarian *r*, it comes as no surprise that almost all subjects fricated and devoiced *v* here, sometimes even *r* itself. Here, too, we ob-served the deletion of *v* with one subject, even when this subject was asked to repeat the test sentence once again. It is thus not unforeseen that assimilation of voicing does not occur: *v* is not actively voiced at all in this position (its is either passively voiced, devoiced or deleted), and so the "Voicing Requirement" is not violated. Figure 7 shows subject 1's pronunciation of *két Wrangler*, in which *v* is realized devoiced/strongly fricated. Clearly, as the lack of glottal pulsing indicates, the cluster [tv̥] is not voiced (actually, even *r* is partially voiceless):
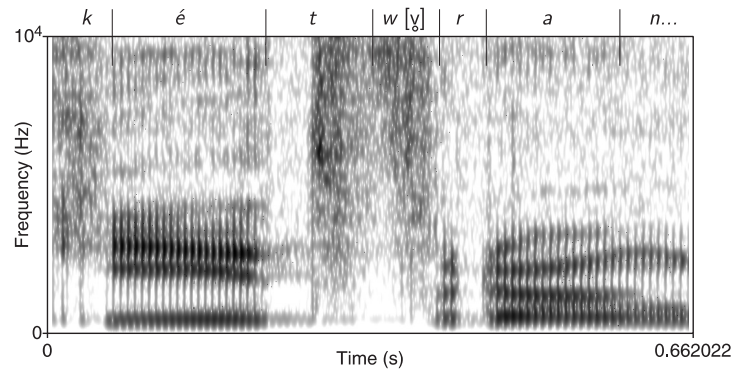
Fig. 7

két Wrangler 'two pairs of Wrangler jeans' (subject 1)

6. Conclusion and remaining issues

The paper has put forth a unified analysis of Hungarian *v*, in which its phonological patterning can be explained in a model based on the phonetic properties of this segment and its linear context. The most important claim has been that the phonetic targets of *v* are inherently contradictory on aerodynamic grounds and can only be maintained in phonetically favourable positions. In such "beneficial" environments, the model predicted the emergence of a passively voiced narrow approximant [ʋ]. Segmental contrast of this consonant has been proposed and found to be most salient and hence preservable between sonorants, vowels and (wide) approximants. In other positions, [ʋ] is predicted to give up one of its phonetic targets. As a result of this, two realizations are possible: when it devoices, it becomes a strongly fricated, noisy sound; when its voicing target is kept, it loses much of its friction. Hungarian has been shown to be a language which prefers the devoicing strategy in unfavourable positions. As a consequence of the phonetic realizations (owing to specific phonetic factors), its phonological behaviour can also be explained: its role in what we called the "Voicing Requirement" as well as its phonotactic distribution. Neither a prevocalic [ʋ], being passively voiced, nor a devoiced [ʋ̥] will induce voicing assimilation as a potential trigger. "Unstable" *v* has been found to be very dependent on the phonetic properties of surrounding sounds. In pre-obstruent position, where it occurs devoiced, active devoicing is easily assimilated by *v*, resulting in a very strongly fricated voiceless sound, which we found very similar

to [f] based on its acoustic parameters. It is thus in this context that the *f–v* contrast is most often lost. If the consonant is voiced after *v*, *v* is realized as a voiced and moderately fricated sound ([v]), this was the case especially before *l*. If, however, a voiced obstruent follows it, *v* is rather devoiced than voiced, therefore, obstruency seems to be a more decisive factor in *v*'s realization than the coarticulation of active voicing. In the surface-oriented model we proposed, thus, all cases of voicing assimilation (including the cases of *lopva/Wrangler*, having been proven to be problematic for formalist analyses) are explained, without exceptions.

The graduality of *v*'s phonotactics is also accounted for in our cue-based approach. In the least saliently cued contexts, it is predicted that *v* does not occur at all or occurs in only a few items. Most tokens are thus to be found in words in which *v* has several cues of good quality that make its contrast better perceivable. Consequently, in the present analysis, words regarded as "exceptional" in formalist/representational models (like *kedv, hamv*) are as much part of the grammar as the "regular" cases; in fact, their **marginal** status in the phonotactic "space" of the grammar is also predicted. The aerodynamics of *v* and its contexts as well as their cueing potential both predict its distributional regularities, they are intimately linked. *v* is devoiced in unfavourable contexts due to aerodynamic reasons, these contexts are the same as those where cues cannot license its contrast either. It is thus in this way that the current model is directly phonetics-based. Phonetic realizations conduct phonological patterning and the realizations themselves are grounded in phonetic factors.

This analysis is admittedly a first step in a phonetically-rooted phonological analysis of Hungarian *v*, as well as its consonantal phonotactics and voicing/voicing neutralization. Therefore, several issues have had to be left untouched here. Further future research is needed to answer such questions as why certain consonants trigger devoicing of *v* to a greater degree than others (cf. the *tv/kv–pv* asymmetry), how exactly the link between final obstruent devoicing and sonorant voicing can be explained in a phonetics-based model, how the neutralization of *v* can be described (is it partial or complete?), if neutralization is partial, what phonetic parameters help maintain/perceive the contrast, how speakers generally perceive and interpret the devoicing of *v*, how the analysis can be extended to other consonants (especially voiced fricatives and approximants) and to other languages, etc.[42] Even though our aim was not to provide a

---

[42] *v* does not behave the same way in all languages. For example, in Polish, voiceless obstruents assimilate to *v* and English *v* (just like the rest of the voiced fricatives)

complete account of how a non-formalist phonological framework is built up, we see the current analysis as a useful contribution to the growing body of work on surface-based functionalist phonology, according to which synchronic grammar (or grammar change)[43] is directly influenced by such low-level functional factors as the aerodynamics of articulation, and speech perception.

Appendix

The following test sentences were used in the experiment:

1. Már megint kiment a kocsimban a bovden.
2. Ennyi bóvlit egy rakáson!
3. A szovjet nagykövet új kocsija egy Chevrolet.
4. Hatvan kilométerre ide van egy medve- és farkas-rezervátum.
5. Mit kotyvasztasz?
6. (a) A likvid tőke egyszerűen azt jelenti, hogy mennyi pénzed van.
   (b) A mezőgazdasági szubvenciókról többet nem akarok hallani, már unom.
7. A szomszéd özvegyasszony állandóan a dugványaival van elfoglalva.
8. Az a fösvény tolvaj a regényben köszvénytől szenved.
9. Nekem nem tetszett Udvaros Dorottya, egyáltalán nem illett hozzá ez a hamvas ártatlan kislány szerep.
10. A sátorponyva alatt hasaltak a fegyveres őrök.
11. Ne a Lacit hívd!
12. Érted? Engem hívj!
13. Ne aggódjatok! Lesz itt jókedv.
14. Már megint volt az asztalán két új könyv.
15. Az angol is germán nyelv.
16. Nemzetközileg melyik a legértékesebb szerv?
17. Magyarország védett ragadozó madarai közül legritkább a sólyom és az ölyv.
18. Vlach János nyugalmazott vezérőrnagy jelentette fel.
19. Wranglert vett magának, nem Trapper farmert.
20. Két Wranglert vett, nem csak egyet.
21. Városunk nagyjairól festmények készülnek.
22. A spanyol konkvisztádorok havonta két várost döntöttek romba.

---

has also been reported to increase the voicing of a preceding obstruent (Jansen 2004). It seems, however, that the functional factors of contrast and contrast dispersion, in addition to the phonetic factors, play an important role in languages where *v* voices: apparently, *v* only triggers voicing assimilation when it is **actively** voiced (like Polish and English *v* and unlike Hungarian *v*). Preliminary data, however, suggest that if a language has a contrastive actively voiced [v], it also has a contrastive passively voiced approximant counterpart, [ʋ/w]. It is as if active voicing plus friction ([v]) are employed so that the contrast can be maximally salient and distinct from passively voiced and weakly fricated or frictionless [ʋ/w]. Such a two- or three-way contrast is reported for a few languages in Ladefoged–Maddieson (1996) as well as Padgett (2002): crucially, the non-approximant [v] in those languages is actively voiced and fricated. The tentative prediction is thus that we should find the Hungarian way of *v*-patterning in those languages in which *v* does not contrast with other voiced labial/labiodental fricatives/approximants. No doubt, much further research is needed in this area, too.

[43] See the diachronic/evolutionary functionalist approach of Ohala (1981; 1993); Blevins (2004), among others.

23. Tviszt, tviszt, tviszt, itt a kaszinó tviszt!

24. Kvarckristályokat találtak a szomszédos hegyekben, azóta olyan sok a látogató.

25. Dvorzsák-szimfóniákat hallgattunk, attól nyugodtunk meg ennyire.

26. Gvárdián, Guatemala, több nem jut eszembe.

27. Cvekedlit? Úgy érted, káposztás cvekedlit kérek-e? Imádom!

28. Szvettert vegyél, hallod? Talán a falnak beszélek?!

29. Svédcseppeket szedtem a torokgyulladásomra, és két nap alatt elmúlt.

30. Veszprém az egyik legszebb város a Dunántúlon.

31. Vasat találtak és rezet, abból gazdagodtak meg.

32. Véletlen egybeesésnek minősítették az esetet, pedig kizárt, hogy csak erről van szó.

33. Vígan ünnepelték a leszerelést.

34. Anélkül, hogy a fejét megmozdította volna, lopva a tigrisre nézett.

35. Vonattal mentünk Avasig, onnan busszal.

36. Vivien, nehogy megkóstold! Az sav!

37. Vulkánkitöréseket észleltek Indonézia partjainál.

38. Vörös zsebkendőt lengetett, ki tudja, mire gondolt.

39. Tegnap felhívtam az orvost.

40. Pannika már nagyon szeretne óvodás lenni.

References

Balise, R. R. – Randy L. Diehl 1994. Some distributional facts about fricatives and a perceptual explanation. In: Phonetica 51 : 99–110.

Barkaï, Malachi – Julia Horvath 1978. Voicing assimilation and the sonority hierarchy: Evidence from Russian, Hebrew and Hungarian. In: Linguistics 212 : 77–88.

Blaho, Sylvia 2005. Another look at the misbehaving segments of Hungarian voicing assimilation. In: Piñón – Siptár (2005, 35–55).

Blevins, Juliette 2004. Evolutionary phonology: The emergence of sound patterns. Blackwell, Oxford & Malden MA.

Boersma, Paul – David Weenink 2005. Praat: Doing phonetics by computer (Version 4.3.19) [Computer program]. Retrieved July 20, 2005. (http://www.praat.org/)

Chomky, Noam – Morris Halle 1968. The sound pattern of English. Harper & Row, New York.

Cyran, Eugeniusz 1997. Resonence elements in phonology. A study in Munster Irish (PASE Studies and Monographs 3). Wydawnictwo Folium, Lublin.

Cyran, Eugeniusz (ed.) 1998. Structure and interpretation. Studies in phonology. (PASE Studies and Monographs 4). Wydawnictwo Folium, Lublin.

Cyran, Eugeniusz – Morgan Nilsson 1998. The Slavic [w > v] shift: A case for phonological strength. In: Cyran (1998, 89–100).

Farnetani, Edda 1997. Coarticulation and connected speech processes. In: William J. Hardcastle – John Laver (eds): The handbook of phonetic sciences, 371–404. Blackwell, Oxford & Malden MA.

Flemming, Edward 2004. Contrast and perceptual distinctiveness. In: Hayes et al. (2004, 232–76).

Fodor, Katalin 2003. A nyelvjárási hangtani jelenségek [Dialectal phonological phenomena]. In: Kiss (2003, 325–50).

Forrest, Karen – Gary Weismer – Paul Milenkovic – Ronald N. Dougall 1988. Statistical analysis of word-initial voiceless obstruents: Preliminary data. In: Journal of the Acoustical Society of America 84 : 115–23.

Fuchs, Susanne – Jana Brunner 2005. Conflicting requirements for voiced alveolar fricatives. Paper presented at the Conference on Turbulences, 13–14 October 2005, Berlin.

Gordon, Matthew – Paul Barthmaier – Kathy Sands 2005. A cross-linguistic study of voiceless fricatives. Ms. University of California, Santa Barbara.

Hall, Daniel Currie 2003. Laryngeal feature specifications in West Slavic languages. In: Toronto Working Papers in Linguistics 20 : 93–114.

Hall, Daniel Currie 2004. A formal approach to /v/: Evidence from Czech and Slovak. In: Olga Arnaudova – Wayles Browne – Maria Luisa Rivero – Danijela Stojanovic (eds): Formal approaches to Slavic linguistics 12: The Ottawa Meeting 2003, 187–206. Michigan Slavic Publications, Ann Arbor.

Halle, Morris 1959. The sound pattern of Russian. Mouton, The Hague.

Hayes, Bruce 1984. The phonetics and phonology of Russian voicing assimilation. In: Mark Aronoff – Robert T. Oehrle (eds): Language sound structure, 318–28. MIT Press, Cambridge MA.

Hayes, Bruce – Robert Kirchner – Donca Steriade (eds) 2004. Phonetically based phonology. Cambridge University Press, Cambridge.

Hume, Elizabeth – Keith Johnson 2001. A model of interplay between speech perception and phonology. In: Elizabeth Hume – Keith Johnson (eds): The role of speech perception in phonology, 3–26. Academic Press, New York.

Imre, Samu 1971. A magyar nyelvjárások rendszere [The system of Hungarian dialects]. Akadémiai Kiadó, Budapest.

Jansen, Wouter 2004. Laryngeal contrast and phonetic voicing: A laboratory phonology approach to English, Hungarian, and Dutch. Doctoral dissertation, Rijksuniversiteit Groningen.

Jassem, Wiktor 1979. Classification of fricative spectra using statistical discriminant functions. In: Björn Lindblom – Sven Öhman (eds): Frontiers of speech communication research, 189–206. Academic Press, New York.

Johnson, Keith 2003. Acoustic and auditory phonetics (Second edition). Blackwell, Malden MA & Oxford.

Juhász, Dezső 2003. A magyar nyelvjárások területi egységei [The regional units of Hungarian dialects]. In: Kiss (2003, 262–324).

Kager, René 1999. Optimality theory. Cambridge University Press, Cambridge.

Kenstowicz, Michael – Mahasen Abu-Mansour – Miklós Törkenczy 2003. Two notes on laryngeal licensing. In: Stefan Ploch (ed.): Living on the edge: 27 papers in honour of Jonathan Kaye, 259–82. Mouton de Gruyter, Berlin & New York.

Kiefer, Ferenc (ed.) 1994. Strukturális magyar nyelvtan 2: Fonológia [A structural grammar of Hungarian 2: Phonology]. Akadémiai Kiadó, Budapest.

Kiss, Jenő (ed.) 2003. Magyar dialektológia [Hungarian dialectology]. Osiris Kiadó, Budapest.

Kiss, Zoltán 2005. Graduality and closedness in consonantal phonotactics: A perceptually grounded approach. In: Sylvia Blaho – Luis Vicente – Erik Schoorlemmer (eds): Proceedings of Console XIII, 171–95. Student Organization of Linguistics in Europe (SOLE), Leiden.

Kiss, Zoltán in preparation. The phonotactics of consonant clusters: A perception and articulation-based approach. Doctoral dissertation, Eötvös Loránd University (ELTE), Budapest.

Kornai, András 1994. On Hungarian morphology (Linguistica, Series A: Studia et Dissertationes 14). Linguistics Institute of Hungarian Academy of Sciences, Budapest.

Ladefoged, Peter 2003. Phonetic data analysis: An introduction to fieldwork and instrumental techniques. Blackwell, Malden MA & Oxford.

Ladefoged, Peter – Ian Maddieson 1996. The sounds of the world's languages. Blackwell, Cambridge MA & Oxford.

Lombardi, Linda 1995. Laryngeal neutralization and syllable wellformedness. In: Natural Language and Linguistic Theory 13 : 39–74.

Lombardi, Linda 1999. Positional faithfulness and voicing assimilation in Optimality Theory. In: Natural Language and Linguistic Theory 17 : 267–302.

Lulich, Steven M. 2004. Russian [v]: An acoustic study. In: Folia Linguistica (Acta Societatis Linguisticae Europaeae) 38 : 63–85.

Machač, Pavel – Pavel Skarnitzl 2005. Spectral moments of Czech plosives. Paper presented at the Conference on Turbulences, 13–14 October 2005, Berlin.

Ohala, John J. 1981. The listener as a source of sound change. In: Carrie S. Masek – Roberta A. Hendrik – Mary Frances Miller (eds): Papers from the parasession on language and behaviour (CLS 17), 178–203. Chicago Linguistics Society, Chicago.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In: P. MacNeilage (ed.): The production of speech, 189–216. Springer, New York.

Ohala, John J. 1993. The phonetics of sound change. In: Charles Jones (ed.): Historical linguistics: Problems and perspectives, 237–78. Longman, London.

Olsson, Magnus 1992. Hungarian phonology and morphology (Travaux de l'Institute de Linguistique de Lund 26). Lund University Press, Lund.

Padgett, Jaye 2002. Russian voicing assimilation, final devoicing, and the problem of [v]. Ms. University of California, Santa Cruz.

Petrova, Olga 2003. Sonorants and the labiodental continuant /v/ in Russian voice assimilation: An OT analysis. In: Olga Arnaudova – Wayles Browne – Maria Luisa Rivero – Danijela Stojanovic (eds): Formal approaches to Slavic linguistics 11: The Amherst Meeting 2002, 413–32. Michigan Slavic Publications, Ann Arbor.

Petrova, Olga – Szilárd Szentgyörgyi 2004. /v/ and voice assimilation in Hungarian and Russian. In: Folia Linguistica (Acta Societatis Linguisticae Europaeae) 38 : 87–116.

Piñón, Christopher – Péter Siptár (eds) 2005. Approaches to Hungarian 9 (Papers from the Düsseldorf conference). Akadémiai Kiadó, Budapest.

Rebrus, Péter – Viktor Trón 2002. A fonotaktikai általánosításokról (Kísérlet a magyar mássalhangzó-kapcsolatok nem-reprezentációs leírására) [On phonotactic generalisations (An attempt at the non-representational description of Hungarian consonant clusters)]. In: Márta Maleczki (ed.): A mai magyar nyelv leírásának újabb módszerei V. Szeged, 2001. szeptember 27–29. [Current methods in the description of Hungarian 5. Szeged, 29–29 September 2001], 17–63. Szegedi Tudományegyetem, Általános Nyelvészeti Tanszék, Magyar Nyelvészeti Tanszék, Szeged.

Ritter, Nancy A. 2000. Hungarian voicing assimilation revisited in Head-Driven Phonology. In: Gábor Alberti – István Kenesei (eds): Approaches to Hungarian 7 (Papers from the Pécs conference), 23–49. JATEPress, Szeged.

Shadle, Christine H. 2005. Characterizing fricatives using dynamics of turbulence noise. Paper presented at the Conference on Turbulences, 13–14 October 2005, Berlin.

Shadle, Christine H. – Sheila J. Mair 1996. Quantifying spectral characteristics of fricatives. Paper presented at the Fourth International Conference on Spoken Language Processing (ICSLP96). Manuscript.

Siptár, Péter 1994. A mássalhangzók [Consonants]. In: Kiefer (1994, 183–272).

Siptár, Péter 1996. A Janus-faced Hungarian consonant. In: The Even Yearbook 2 : 83–96.

Siptár, Péter – Szilárd Szentgyörgyi 2002. *H* as in Hungarian. In: Acta Linguistica Hungarica 49 : 427–56.

Siptár, Péter – Miklós Törkenczy 2000. The phonology of Hungarian. Oxford University Press, Oxford.

Steriade, Donca 1997. Phonetics in phonology: The case of laryngeal neutralization. Manuscript. UCLA.

Steriade, Donca 1999. Alternatives to syllable-based accounts of consonantal phonotactics. In: O. Fujimora – B. Joseph – B. Palek (eds): Proceedings of the 1998 Linguistics and Phonetics Conference, 205–42. The Karolinum Press, Prague.

Stevens, Kenneth N. 1998. Acoustic phonetics. MIT Press, Cambridge MA.

Szentgyörgyi, Szilárd – Péter Siptár 2005. Hungarian *H*-type segments in Optimality Theory. In: Piñón – Siptár (2005, 261–81).

Szigetvári, Péter 1998a. Voice assimilation in Hungarian: The hitches. In: The Even Yearbook 3 : 223–36.

Szigetvári, Péter 1998b. Why [h] is not voiced. In: Cyran (1998, 287–301).

Szépe, György 1968. Az alsóbb nyelvi szintek leírása [The description of low linguistic levels]. In: Általános Nyelvészeti Tanulmányok 6 : 359–466.

Törkenczy, Miklós 1994. A szótag [The syllable]. In: Kiefer (1994, 273–392).

Törkenczy, Miklós 2000. Lombardi (1999) on voicing assimilation in final clusters in Yiddish-type languages: The embarrassing case of Hungarian. Paper presented at the Massachusetts Institute of Technology, Cambridge, Massachusetts.

Vago, Robert 1980. The sound pattern of Hungarian. Washington D. C., Georgetown University Press.

Wright, Richard 2004. A review of perceptual cues and cue robustness. In: Hayes et al. (2004, 34–57).

Zsigri, Gyula 1994. Magyar mássalhangzószabályok [Rules for Hungarian consonants]. Doctoral dissertation, József Attila University (JATE), Szeged.