THE REALISATION OF VOICING ASSIMILATION RULES IN HUNGARIAN SPONTANEOUS AND READ SPEECH: CASE STUDIES*

ALEXANDRA MARKÓ[a] – TEKLA ETELKA GRÁCZI[b] – JUDIT BÓNA[a]

[a]Department of Phonetics, Eötvös Loránd University
Múzeum krt. 4/A, H–1088 Budapest, Hungary
marko.alexandra@btk.elte.hu; bona.judit@btk.elte.hu

[b]Research Institute for Linguistics, Hungarian Academy of Sciences
Benczúr u. 33., H–1068 Budapest, Hungary
graczi@nytud.hu

Abstract: Hungarian represents a particularly fruitful ground for exploring voicing assimilation. Although this topic has been extensively analysed, a contradiction can be observed between most phonological descriptions and acoustic-phonetics-based studies of voicing assimilation. Theoretical works suggest that this process in Hungarian speech is a purely regressive, obligatory and categorical phenomenon, but in practice divergent realisations can be observed. In the present paper three case studies of voicing assimilation are performed. CCC clusters, CC clusters interrupted by pause and partially voiced realisations were analysed. The results showed that in the first two cases the speech planning process and the degree of self-monitoring were the most influential factors, while the various concomitances of voicing and devoicing arising due to aerodynamic and articulatory reasons resulted in partially voiced realisations. The variability of the data confirms the hypothesis that Hungarian voicing assimilation is a gradient and sometimes only partly regressive process. Even if it operates mainly obligatorily, several factors can override it.

Keywords: Hungarian voicing assimilation, pauses, partially voiced consonants, three-member clusters, speech planning

1. Introduction

In the past few decades numerous studies have examined phonetic evidence for phonological processes (e.g., Barry 1992; Recasens–Pallarès

---

2001; Jansen 2007). An increasing number of phonetic studies have been focusing on neutralisation due to voicing assimilation, a phonological process by which one sound assumes the voicing feature of a neighbouring sound (Burton–Robblee 1997; Jansen–Toft 2002; Gow–Im 2004, etc.). In phonetic investigations of voicing assimilation, it is not only the question of whether it neutralises voicing contrasts that arises, but also that of the nature, direction and degree of assimilation. Hungarian offers particularly rich opportunities for exploring voicing assimilation.

Differences in speech styles result in differences in phonetic characteristics, both segmental and suprasegmental. Several models of coarticulation take speech styles into consideration as the source of one of the prime effects on realisations. In his H&H model, Lindblom (1990) discusses two modes of speech, hyper- and hypo-speech. He suggests that a sound (hence, a sequence of sounds) is never pronounced twice in the same way, for speech always varies along several continua such as context, loudness, rate of speech, speaking style and so on. Lindblom proposed that speakers continually monitor how clear their articulation should be in view of the information assumed to be shared by the speaker and the listener. We assume that spontaneous and read speech represent two different intervals of this H&H scale: reading aloud is usually located closer to maximally carefully articulated hyper-speech, while spontaneous speech variants are typical manifestations of hypo-speech. Therefore, dissimilarities can be observed in the phonetic characteristics of sounds (coarticulation, reduction, etc.). Also, these diverse speech styles involve dissimilarities in speech planning. In read speech, the speaker need not be concerned about higher level planning processes; he/she can spend more effort on articulation. As Levelt (1989, 259) wrote: "In reading, the speaker can rely heavily on the printed materials. Lexical retrieval and the building of syntactic constituents can be based largely on parsing of the visual input. Reading out is primarily a perceptual, phonological and articulatory task." It does not require message planning and grammatical encoding. Spontaneous speech production, on the other hand, includes several consecutive and parallel cognitive processes; therefore articulation is necessarily far more automatic. In Levelt's (1989) model, "grammatical encoding, form encoding, and articulating [. . .] are assumed to be automatic to a large degree. They are speedy and reflex-like, require very little attention, and can proceed in parallel" (*op.cit.*, 28). Our hypothesis is therefore that the above-mentioned differences lead to dissimilarities in voicing assimilation between reading and spontaneous speech.

In Hungarian, adjacent obstruents mostly agree in terms of voicing, e.g., *bukta* [buktɔ] 'sweet roll', *kesztyű* [kɛscyː] 'glove'; *labda* [lɔbdɔ] 'ball', *pezsgő* [pɛʒgøː] 'champagne'. Loanwords that originally contain an obstruent cluster of heterogeneous voicing automatically get adjusted to this pattern: *futball* [fudbɔl] 'football', *abház* [ɔphaːz] 'Abkhaz'.

In suffixed forms, stem-final voiceless obstruents become voiced if the suffix begins with a voiced obstruent (1) and vice versa: stem-final voiced obstruents become voiceless if the suffix begins with a voiceless obstruent (2). The process also applies across a compound boundary, across a word boundary, and (theoretically) across any higher prosodic boundary as long as no pause intervenes (Siptár–Törkenczy 2000).

(1)   Voicing

zsák [ʒaːk] 'sack' + ban [bɔn] 'inessive' → zsákban [ʒaːgbɔn] 'in (a) sack'

kis [kiʃ] 'small' + béka [beːkɔ] 'frog' → kis béka [kiʒbeːkɔ] 'small frog'

(2)   Devoicing

kéz [keːz] 'hand' + től [tøːl] 'ablative' → kéztől [keːstøːl] 'from (a) hand'

nagy [nɔɟ] 'big' + kutya [kucɔ] 'dog' → nagy kutya [nɔckucɔ] 'big dog'

This regressive assimilation applies iteratively to affect more than one obstruent in a cluster: e.g., *test-ben* [tɛʒdbɛn] 'body-inessive' (Kenesei et al. 1998). Hungarian voicing assimilation is considered to be a postlexical, obligatory phonological rule that does not depend on speakers' articulation rate. Sonorants do not participate in the process (Siptár–Törkenczy 2000).

There are three segments in Hungarian that behave asymmetrically with respect to voicing assimilation. /v/ undergoes devoicing (*szívtől* [siːftøːl] 'heart-ablative') but does not trigger voicing (*hatvan* [hɔtvɔn] 'sixty'). (For the "double-faced" phonological behaviour of /v/ see Kiss–Bárkányi 2006.) On the contrary, /h/ (or /x/ according to the theoretical approach of Siptár–Törkenczy 2000) triggers devoicing (*adhat* [ɔthɔt] 'he may give') but does not undergo voicing before an obstruent. Sonorant /j/ is realised as a palatal approximant in most cases (*ajtó* [ɔjtoː] 'door'). However, when preceded by a consonant and followed by a pause, it surfaces as a fricative, which is voiceless if preceded by a voiceless obstruent (*kapj* [kɔpç] 'get-imperative') and voiced after voiced obstruents (*dobj* [dobj] 'throw-imperative') and sonorants (*férj* [feːrj] 'husband').

This means that when /j/ appears as an obstruent, it does not trigger regressive voicing assimilation, but it undergoes progressive voicing assimilation when preceded by an obstruent (Blaho 2008).

Several phonological accounts of Hungarian voicing assimilation can be found in the literature—besides the papers referred to above. Vago (1980) gave a complete description of Hungarian phonological processes (in an SPE-type binary feature model). Szigetvári's (1998) study dealt with the status of the problematic /h/, /v/ and /j/, while Ritter (2000) applied a head-driven phonology approach to them. Jansen (2004) carried out a functional laboratory phonological analysis primarily in terms of the tenseness of the participating obstruents. His results show that Hungarian regressive voicing assimilation leads to incomplete neutralisation of [tense] distinctions in target sounds. Studies based on acoustic analyses of Hungarian voicing assimilation data are also consistent with each other in considering the nature of an 'intermediate' output of the assimilation process. Jansen and Toft (2002) argue that Hungarian voicing assimilation may be gradual. The analysis by Gow and Im (2004) yielded the result that Hungarian fricative voicing assimilation produces segments whose voicing is acoustically between those of voiced and voiceless fricatives. Kiss and Bárkányi (2006) proposed a unified, surface-based functionalist analysis of the phonology of Hungarian *v*. Their analysis was grounded in the aerodynamics of *v*'s articulation as well as in the relative perceptibility of its contrast in various contexts.

The majority of earlier investigations were carried out on read material: single words, two-word sequences or sentences. Voicing assimilation in Hungarian spontaneous speech has been examined by Gósy (1999; 2002). Her studies focused on the voicing assimilation process at word boundaries in the case of pauses (in connection with activation processes of the mental lexicon). She found that if the pause between two words does not exceed 55 ms, voicing assimilation always takes place. If the pause exceeds 314 ms, the process does not apply. Between these two values the speech production mechanism is either able to activate voicing assimilation rules together with selecting the words, or fails to do so.

A contradiction can be observed between most phonological descriptions and acoustic-phonetics-based studies of voicing assimilation. Theoretical works (except for Vago 1980) suggest that this process is a purely regressive, obligatory and categorical phenomenon in Hungarian speech, regardless of juncture strength as long as no physical pause intervenes. However, in actual fact, divergent realisations can be observed (see the sources above).

In our former study (Bóna et al. 2008) three speaking styles (reading, interpreting and spontaneous) were compared with respect to the realisation of the voicing assimilation process. 10 Hungarian native speakers participated in the experiment. One of their tasks was to read aloud a 13-sentence-long newspaper text and 14 independent sentences. The other speech sample type was quasi-spontaneous or interpreted speech in which the participants had to relate the contents of a recorded text they had listened to. The third recording was that of spontaneous speech from discussions on various topics. All CC clusters (and only the two-member ones) both within words ($C_1 C_2$) and across word boundaries ($C_1 \# C_2$) that were not interrupted by pauses of any length and consisted of voiced and voiceless obstruents (in any order) were examined acoustically. In 88.5% of the 1,190 CC clusters analysed, the process of voicing assimilation operated regularly. The irregular cases (11.5%) could be divided into four subgroups:

  (i) The assimilation process did not operate at all in 3.1% of the corpus.
 (ii) Voicing assimilation was progressive in 1.3% of the cases.
(iii) The output was partially voiced in 4.5% of the data.
(iv) C_1 deletion occurred in 2.6% of the cases.

The ratio of voicing and devoicing showed no difference: voicing took place regularly in 85.8% and devoicing in 89.3% of the relevant tokens. According to the results in terms of speech styles, major differences could be observed between read and spontaneous (including interpreted) speech.

From this work several questions have arisen that further analyses were designed to answer. The earlier study did not take into account CCC clusters in which two adjacent obstruents differed in voicing. In the present paper, we summarise the results of an analysis of such sequences. Also, CC clusters of two adjacent obstruents interrupted by pauses of any length (either within words or across word boundaries) were excluded from the former material. These are also examined in this paper, in which we interpret our results in comparison to Gósy's (1999) data.

According to our former study, a partially voiced output was the most frequent realisation among the irregular ones. We found a great deal of inter-speaker variability in partially voiced realisations, which could be attributed to individual articulation properties. The present study analyses this type of voicing assimilation output on the basis of a much larger number of samples. The data confirm that Hungarian voicing assimilation is a gradient phenomenon.

2. Material and method

The material of the present experiment was selected from the BEA Hungarian speech database ("BEA" stands for ***be****szélt nyelvi **a***datbázis* 'spoken language data base'), recorded at the Research Institute for Linguistics of the Hungarian Academy of Sciences under constant circumstances in an anechoic chamber (http://www.nytud.hu/dbases/bea/index.html; for detailed technical parameters, see Gósy 2008). BEA's material consists of several speech samples from various types of spontaneous speech, repetitions of stimuli, reading aloud, and conversations. The subjects are monolingual native speakers of standard Hungarian aged between 20 and 70 years.

For the present research, ten speakers' material was selected. The five female and five male subjects were selected according to age and level of education. Their age ranged from 29 to 64 years. None of the participants had discernible uncorrected deficits in visual perception, speaking or hearing, nor did they have known reading difficulties. Reading and two types of spontaneous speech samples were used for analysis as follows.

One of the subjects' tasks was reading aloud a 13-sentence-long newspaper text and 25 independent sentences, a total of 433-word speech material by speaker. The read subcorpus of our analysis consisted of these recordings. The time of reading aloud varied between 209.6 and 300.0 seconds, adding up to a total of 43.3 minutes.

The other speech sample types were quasi-spontaneous (interpreted) and spontaneous speech. In the first case, the participants had to relate the contents of a recorded text they had listened to, while in the second part an interview was carried out on several topics (basically on the subject's job and hobbies). On the basis of the results of our earlier research on voicing assimilation in spontaneous speech, these spontaneous speech samples were handled together in the spontaneous subcorpus. The length of these samples was between 6.7 and 11.3 minutes, and they contained 638 to 1337 words. This subcorpus of 84.3 minutes of speech material consists of 10,717 words.

The material was analysed from three different points of view:

(i) CCC clusters in which two adjacent obstruents differ in voicing were analysed in terms of the order of the consonants, the position of clusters (within a word or across a word boundary), and the type of the third consonant (a sonorant or the obstruent *v* that operates irregularly in terms of voicing assimilation). The output of the voicing assimilation

process was categorised, regarding also possible changes of the third consonant. Clusters that were interrupted by a pause were excluded from the material of this examination.

(ii) For C_1C_2 obstruent clusters that were interrupted by pauses of any length (either within words or across word boundaries), the type (silent, filled or combined) and duration (in milliseconds) of pauses were defined. The output of the voicing assimilation process was analysed in terms of whether it was realised regularly or not. It was also determined where (between what types/parts of words) the pause was found and what the reason of its appearance could be. For the possible explanation(s) we checked the narrow context including disfluency phenomena and always took the speaker's own articulation characteristics or habits into consideration. When C_2 was voiceless, a pause between C_1 and C_2 was measured as such only if it exceeded 100 milliseconds to make sure that an occlusion phase was not considered to be a pause. For the same reason, the last 50-millisecond part of a "pause" preceding a stop release was not taken into account in the measurements.

(iii) Partially voiced output of voicing assimilation: we analysed $C_1(\#)C_2$ obstruent clusters not interrupted by a pause and realised ambiguously in terms of the voicing feature. The ratio of the voiced part was determined, and the tokens of voicing assimilation where the output was partially voiced were classified. Clusters that were interrupted by a pause or were combined with a third consonant were excluded from the material of this examination.

Voicing character was determined on the basis of subjective and objective evaluation of both spectrographic and oscillographic data. The presence of the voicing part on the spectrogram and the pulses detected by Praat were considered. Measurements were carried out manually. A consonant was considered as voiced if it contained a quasi-periodic signal in at least 80% of its duration. A consonant was considered as voiceless if it contained quasi-periodic signal in at most 20% of its duration. Between these values the consonant was evaluated as partially voiced (cf. Figure 1). The position of the voiced part(s) was not taken into consideration. The present study analyses the acoustic presence of vocal fold vibration on the basis of waveforms and spectrograms, and does not take other acoustic cues for voicing like duration of the consonants or that of the preceding vowel into consideration. However, these properties could raise further questions (cf., e.g., Jansen 2004), and are to be analysed in further investigations.
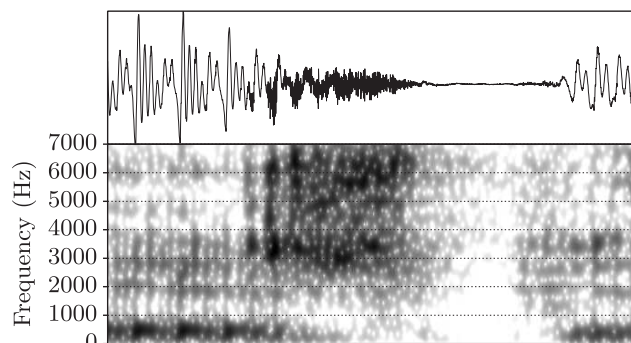
Fig. 1

Oscillogram and spectrogram of the sequence [ɛʒpi] from the words
*telj**es bi**zonyossággal* 'with absolute certainty'

Samples in which creaky voice or any kind of "noise" (laugh, back-channel signal) was detected were excluded from the analysis. The results were analysed in terms of speech styles to see whether they differ between reading and spontaneous speech. For the acoustic analysis, Praat 5.0 (Boersma–Weenink 2005) was used and statistical analyses (independent samples t-test and one-way ANOVA) were carried out using SPSS 16.0.

3. Results

3.1. CCC clusters

The written text contains five CCC clusters, meaning that a total of 50 occurrences were realised by the ten subjects in the read subcorpus. These clusters contain sonorants (*n, r, j*) or (irregular) *v* as third consonant. The cluster $C_{obstr1}C_{obstr2}$ was preceded by the third consonant in two cases (*szervezetü**nkb**e* 'into our organism', *hazá**nkb**an* 'in our country'), while in three cases the third consonant followed the $C_{obstr1}C_{obstr2}$ cluster (*súlyosa**bb pr**oblémát* 'more severe problem-acc', *visszamara**dt v**egyi* 'residual chemical [material]', *a**zt j**elenti* 'it means'). Therefore, the ratio of expected voicing and devoicing and the ratio of realisations appearing across a word boundary and within a word were the same (40 vs. 60%).

C_{obstr1} was deleted in 13 tokens (26.0%) of the read data: *szerveze-tü**nb**e, hazá**nb**an.* In these occurrences it cannot be established whether voicing assimilation took place or not because the potentially assimilated sound cannot be detected acoustically.

In 37 cases (74.0%) voicing assimilation took place regularly. In two of them the third consonant cannot be detected: [hɔzaːgbɔn]. In three further tokens voicing assimilation did occur at the phonological level but C_{obstr2} did not appear on the surface: [ɔsjɛlɛnti]. In one case the assimilation operated (the result was voiceless in accordance with the rule) but the output was different from the expectations regarding the manner of articulation of the target sound: the process resulted in [ɔst͡sjɛlɛnti] instead of [ɔst jɛlɛnti].

78 clusters were analysed in the spontaneous speech material. Voicing was expected in 12 tokens (15.4%), while devoicing would be regular in 66 occurrences (84.6%). Across word boundaries 57 realisations (73.1%) were found, within words 21 (26.9%) occurred.

The CCC clusters of spontaneous speech can be divided into two main groups. In the first one, the third consonant is not expected to influence the voicing assimilation process between C_{obstr1} and C_{obstr2}, since it is a sonorant or *v* that does not trigger voicing. In the second group all the three consonants of the clusters are obstruents (excluding *v*), where regressive assimilation applies iteratively.

The third consonant was a sonorant (*m, n, j, r*) or *v* in 60 cases (76.9% of all spontaneous tokens). Considering these 60 cases to be 100%, in 56.7% of them (34 cases) the $C_{obstr1}C_{obstr2}$ cluster was preceded by the third consonant: e.g., *ma**jd p**edig* 'and then', *har**cb**an* 'in fight'. In 43.3% (26 cases), the third consonant followed the $C_{obstr1}C_{obstr2}$ cluster, e.g., *a**zt j**elentette* 'it meant', *me**gpr**óbál* '(he) tries'.

Among these cases three tokens (5.0%) were observed where C_{obstr1} was deleted, therefore it could not be determined whether or not voicing assimilation took place: e.g., *ba**nkb**an* 'in a bank' realised as [bɔmbɔn], *vo**lt gy**erekkoromban* [vol ɟɛrɛkːorombɔn] '(there) was in my childhood'. These tokens can be considered outputs of fast cluster simplification ("an optional deletion process that targets consonants flanked by consonants, i.e., it deletes the middle one of a sequence of three consonants" in fast speech; cf. Siptár–Törkenczy 2000, 293).

49 occurrences (81.7%) were regular in terms of voicing assimilation. All three consonants were realised as expected in 38 tokens (77.6% of the 49 occurrences). Voicing assimilation applied regularly but C_{son} was deleted from *pé**nzt**ár* [peːstaːr] 'cash-desk'. In 9 cases (18.4%) C_{obstr2} was deleted, while C_{obstr1} became voiceless; e.g., *e**zt n**agyon* [ɛs nɔjon] 'this-acc very much', *a**zt m**egfizetni* [ɔs mɛkfizɛtni] 'to pay for that', *a**zt l**átod* [ɔs laːtod] 'you can see that'. One more occurrence was found where

the output was different from the expectations regarding the manner of articulation of the target sound: *harcban* would be expected to be pronounced as [hɔr$\widehat{d}$zbɔn], but it was realised as [hɔrzbɔn].

Examples of the 8 (13.3%) irregular occurrences: *hangszer* realised as [hɔŋgsɛr] 'instrument', *mint dolgozó* [mint dolgozoː] 'as a worker'. In one of these cases, C_{obstr2} disappeared before voicing assimilation had taken place: *az francia* [ɔz rɔn$\widehat{ts}$iʲɔ] 'it is French'.

The second main group—where all the three consonants of the clusters are obstruents (excluding *v*)—contains 18 tokens (23.1% of all CCC clusters of the spontaneous subcorpus). In the occurrences of this group voicing assimilation applied mostly regularly (16 cases, i.e., 88.9%). Seven cases fulfilled the phonological expectations: $C_{obstr1}C_{obstr2}C_{obstr3}$ was realised, e.g., as *azt hogy* [ɔst hoɟ] 'it-acc that', *azt hitte* [ɔst hitːɛ] '(he) believed that'.

In eight cases C_{obstr2} did not appear on the surface, while voicing assimilation took place, e.g., *azt fogják* [ɔs fogjaːk] 'they will', *azt hogy* [ɔs hoɟ]. Two courses of events can be assumed in the background of these realisations. One possibility is that voicing assimilation progressed iteratively (from C_{obstr3}) and in a later phase C_{obstr2} disappeared: /zth/ > /sth/ > [sh]. As the other way leading to this output, it can be supposed that first fast cluster simplification deleted C_{obstr2}, then C_{obstr3} triggered the assimilation of the voicing feature of C_{obstr1}: /zth/ > /zh/ > [sh]. In one token of this subgroup, one more step was taken: *azt hiszem* 'I think' realised as [ɔsisɛm] with deletion of the *h,* which can be explained by the very frequent occurrence of this phrase as a discourse marker. C_{obstr3} also disappeared in another token: *azt hogy* [ɔst oɟ].

The two remaining occurrences are: *azt gondolom* 'I think' realised as [ɔs gondolom], *azt befejeztem* [ɔs bɛfɛjɛstɛm] 'I have finished it', where iterative voicing triggered by C_{obstr3} was expected, yet voicing assimilation applied only between C_{obstr2} and C_{obstr1}, and then C_{obstr2} was deleted. Therefore these clusters are irregular in terms of voicing assimilation.

3.2. CC clusters interrupted by pause

Both read and spontaneous subcorpora contained C_1C_2 clusters consisting of voiced and voiceless obstruents (in any order) that were interrupted by a pause. In reading, 25 occurrences could be found; in the spontaneous material 64. It is obvious that different reasons can be found in the

background of this phenomenon in the case of reading and spontaneous speech because of the dissimilar speech planning processes. Therefore, in this case study, the separate analysis of C_1C_2 clusters of the two speech types is particularly important.

In reading aloud, voicing should have operated in 22 cases, while only 3 sites for devoicing were found. The length of pauses interrupting C_1C_2 clusters varied between 19 and 714 ms. All the pauses were silent as expected. The voicing assimilation process did not operate at all, with the single exception of a pause realised in 277 ms. This was the only case in the material where the pause occurred within a word, between a stem and an affix (*szerek**b***ől* [sɛrɛg | bøːl] 'from materials'). The pause could be caused by a slip of the tongue or reading uncertainty, where voicing was encoded successfully from the visual information of *b* but the reader was not certain in the cont⸻⸻ The ratio ⸻⸻ of C_1[pause]C_2 realisations in reading ⸻⸻
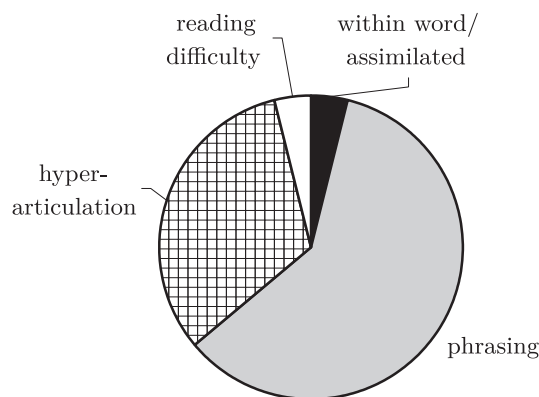

Fig. 2

The ratio of categories of C_1[pause]C_2 realisations in reading aloud

In 15 cases out of the 24 where the assimilation did not operate, the pause occurred between words which were separated by a comma in the written text (being parts of an enumeration): *zavaroka**t**, **d**aganatos* [*megbete-gedéseket*] 'problems-acc, [cases-acc] of tumour'; *zöldsége**k**, **gy**ümölcsök* 'vegetables, fruits'. In these occurrences, phrasing overrode the operation of the phonological process. The length of these pauses varied between 45 and 714 ms; the mean was 286.3 ms, SD was 148.5 ms.

It is unambiguously the case that another eight instances of non-assimilation were caused by hyperarticulation: *amelyeke**t** **gy**anútlanul*

'which-acc unsuspecting', *gyereke**k b**ukfencezni* 'children to tumble', *kala-uz **sz**igorúan* 'controller strictly'. Hyperarticulation was extremely strong in one item where the female speaker inserted a schwa between C_1 and C_2: *kalauzə* [pause 47 ms] *szigorúan*. Pause duration ranged from 19 to 120 ms. The average was 71.6 ms, SD was 26.5 ms. The peculiar differences of pause duration between phrasing and hyperarticulation (despite of some overlapping) can be seen in Figure 3.

One case where a 497 ms pause was realised in *tulipánágyáso**k gy**omlálásával* 'by weeding of tulip beds' probably originated in the 64-year old speaker's temporary reading difficulty.
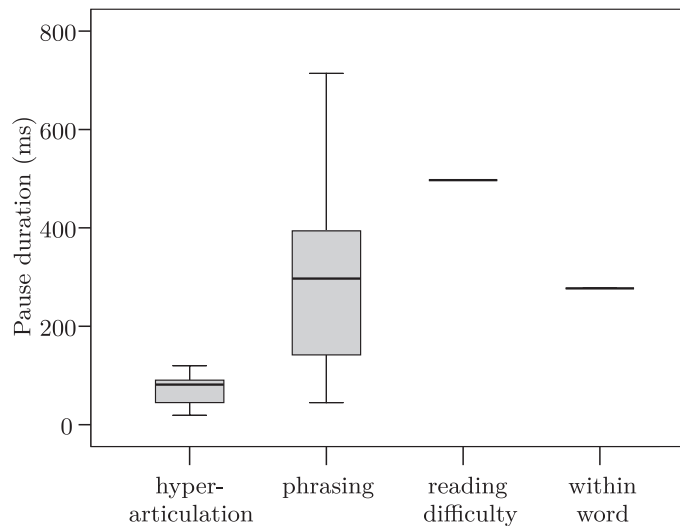

Fig. 3

The range of pause duration between C_1 and C_2 in reading: between words
(hyperarticulation, phrasing and reading difficulty) and within word

In BEA's written texts the number of potential CC voicing assimilation sites is 48, of which 15 crosses word boundaries (8 in the isolated sentences, 7 in the coherent text). The data show that speakers mostly apply the phonological rule in these cases without pausing. If we compare the 24 actual occurrences to the 150 potential sites of the 10 speakers, this means that only 16.0% of $C_1\#C_2$ clusters were realised with a pause. Nevertheless, in all the cases where pause occurred at a word boundary, it blocked the operation of the voicing assimilation process independently of the duration of the pause. Considering the subjects one by one, this

happened at least once but at most four times in any given speaker's read speech material.

All of the 64 C_1[pause]C_2 data of the spontaneous subcorpus occurred at word boundaries. There were 1.3 sites for voicing assimilation per minute in general across word boundaries interrupted by a pause of any length. This means that almost every fifth $C_1 \# C_2$ cluster was interrupted by a pause.

At C_1[pause]C_2 sites, voicing assimilation worked in 11 cases (17.2%), while it did not take place in 53 cases (82.8%). The number of possible devoicing significantly exceeds the number of possible voicing. 47 clusters were found in which dev[...] [v]oicing assimilation took place in only [...] al 17 clusters were annotated where [...] were realised with non-application of [...] he pause that appeared between C_1 a[nd ...] C_1[pause]C_2 realisations in spontaneo[us spe]ech can be see[n ...]
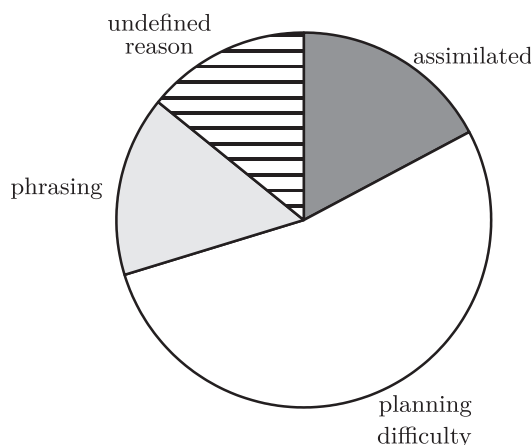

Fig. 4

The ratio of categories of C_1[pause]C_2 realisations in spontaneous speech

Pause duration in the occurrences of regular devoicing ranged between 105 and 767 ms, the average was 328.8 ms, SD was 146.9 ms. In several cases of this type of realisations, the question arises whether the voiceless output really originates from the regular voicing assimilation process. Consider the token with the longest pause: the speaker repeated the conjunction 'that' (*hogy ... hogy* [hoc | hoɟ]), and stopped between the two occurrences gaining some time for speech planning. Repetition and pause

are both considered disfluency phenomena that help the speaker find out how to continue the utterance. The same strategy can be observed in the example where the second longest pause (501 ms) appeared. Between the words *hogy ... felismerik* [hoc | fɛliʃmɛrik] 'that (he) will be recognised' a combined pause can be detected: a 285 ms silent pause followed by a 216 ms schwa hesitation (filled pause). While silent pause is multifunctional (it can be used in a rhetoric function, or to allow time for breathing, but also to help the planning process, self-monitoring etc.), hesitation is generally considered a sign of disharmony in terms of speech planning (cf. Gósy 2005). So the following question arises: If the speaker does not know how to continue the utterance, why would he/she apply a regressive assimilation rule? Therefore it can be assumed that in fact it was not a voicing assimilation process but phrase/word-final devoicing that took place in these (and some other?) cases.

Word-final devoicing (a process by which contrasts between voiced and voiceless obstruents are neutralised in word-final position in favour of the voiceless member) is not considered operative in the Hungarian phonological system (see, e.g., Szigetvári 1998); however, it can be an individual articulation habit (as a consequence of certain aerodynamic effects). Our data are rather diverse so neither voicing assimilation nor word-final devoicing can be unambiguously proven case by case. Nevertheless, three facts should be noted. First, regular voicing was not found among the C_1[pause]C_2 data at all. Second, the 11 regular devoicing occurrences appeared in only 5 speakers' records, and 5 occurrences were documented in the same subject's speech. Finally, in our earlier study, for the 1,190 C_1C_2 realisations of the whole corpus (including $C_1\#C_2$ occurrences but excluding C_1[pause]C_2 ones) the ratio of voicing and devoicing showed no difference: voicing took place regularly in 85.8% and devoicing in 89.3% of the tokens (Bóna et al. 2008). Hence it is not likely that voicing assimilation alone explains the asymmetry of the data in the present case.

In cases where the voicing assimilation process was cancelled (independently of the voicing/devoicing direction), the pauses ranged from 57 to 2245 ms, their mean was 585.2 ms, SD was 321.7 ms. When voicing assimilation took place, the average of pause duration was 328.8 ms (SD was 146.9 ms, data ranged between 105 and 767 ms). Figure 5 shows the distribution of pause durations in two types of realisations: where voicing assimilation operated and where did not. The independent samples t-test shows significant difference between these groups: $t(62) = -1.991$; $p = 0.051$.
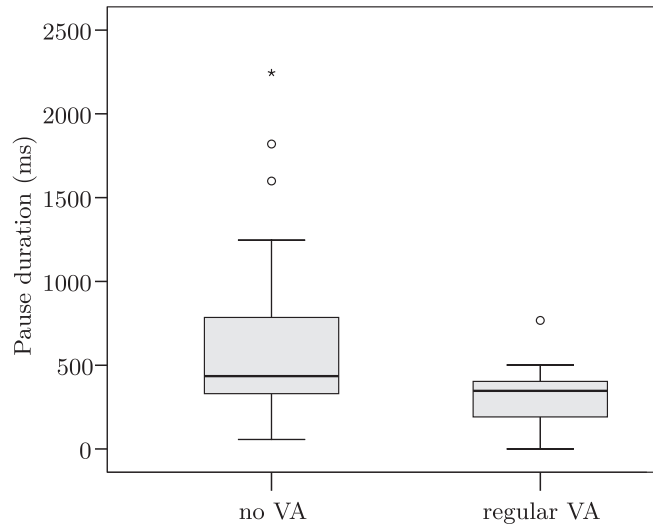
Fig. 5

The range of pause durations between C_1 and C_2 in spontaneous speech:
regular and cancelled voicing assimilation

A planning difficulty was supposed to be the reason of cancelled voicing assimilation in 34 cases (64.2%) out of 53 due to some type of disfluency phenomenon. In 25 occurrences, schwa hesitation (21 times combined with silent pause), 5 times filler words, twice vowel lengthening, once a restart and once an incomplete sentence were documented. (Of course, planning difficulties could be assumed in the background of other occurrences as well, but no disfluency phenomena confirmed this assumption in other cases.) The pauses associated with this type of realisations ranged between 146 and 2245 ms, 714.7 ms on average, where SD was 359.9 ms.

In 10 further cases (18.9%) phrasing resulted in a pause, overriding the voicing assimilation rule. Pauses were shorter in these occurrences: they varied between 57 and 703 ms, their mean was 364.9 ms and SD was 156.4 ms.

On the basis of the surface structure of the remaining 9 (17.0%) $C_1[\text{pause}]C_2$ tokens and their context, it is not clear what type of process cancelled voicing assimilation there. Of course, it could be supposed that some speech planning uncertainty resulted in these realisations even if its traces cannot be observed on the surface. At the same time, these pauses are shorter than those unambiguously associated with planning difficulties. Their duration was measured as 98 to 730 ms, 341.0 ms on average, and with 160 ms SD. Nevertheless, in the earlier study, in the total corpus

of 1,190 obstruent clusters analysed, 37 realisations (3.1%) were found in which voicing assimilation did not take place at all, although the sequence of obstruents was not interrupted by pause. This means that the cancellation of voicing assimilation is not necessarily caused by a pause. It should also be noted that 5 of these 9 occurrences appeared in the same speaker's production, who also relatively often cancelled voicing assimilation in C_1C_2 clusters without pausing.

Figure 6 shows that pauses caused by phrasing are remarkably shorter than those originating in some kind of planning difficulty. According to the boxplot, realisations that could not be connected to a certain process are similar to the phrasing data. ANOVA results show that the variance of the groups is significant: $F(2, 50) = 4.785$; $p = 0.013$.
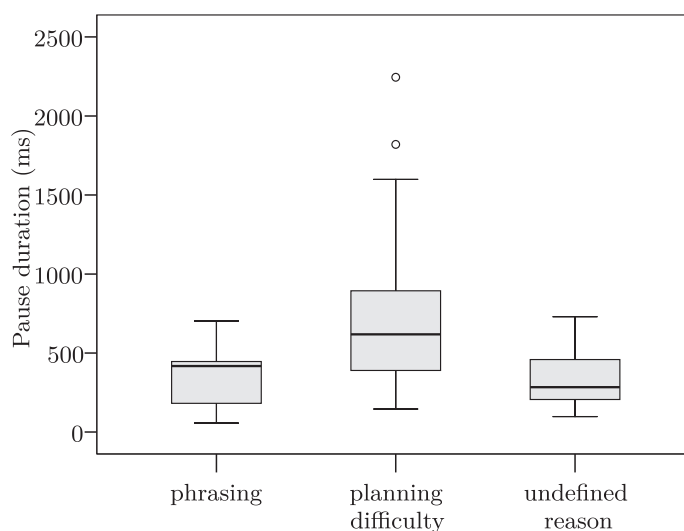

Fig. 6

The range of pause durations between C_1 and C_2 in cases of cancelled voicing assimilation in spontaneous speech: phrasing, planning difficulty, and undefined reason

3.3. Partially voiced output of voicing assimilation

346 realisations were found (91 in reading and 255 in spontaneous speech) where at least one of the occurring sounds was partially voiced. During the analysis of the data, a large difference has been experienced between

cases in which voicing would be the regular process, and those ones in which devoicing would be expected.

The distribution of the 91 partially voiced tokens in the reading data was as follows: 51.6% (47) of the cases were expected as devoicing and 48.4% (44) of the occurrences were expected as voicing. The ratio of between- and within-word clusters was approximately equal. 55.3% (25 cases) of the devoicing and 55.8% (26 cases) of the voicing sites were found in within-word clusters, and the remaining 44.7% (21 cases) of the devoicing sites and 44.2% (20 cases) of the voicing sites were found in word boundary clusters. Among the partially voiced clusters in the reading task, six types of realisations were found that can be classified into three major groups.

The results of the devoicing sites in the read speech samples showed the following distribution. In the first major group, C_1 was partially voiced—this type occurred in 37 cases (78.7% of the tokens). In all of these cases C_2 was unchanged, e.g., *ügyfeleknek* [yɟfɛlɛknɛk] 'for clients' (1a in Figure 8, p. 233). This means that voicing assimilation operated partially on C_1.

In the second major group of devoicing, C_2 was partially voiced. 6 tokens (12.8% of the cases) were realised in this way. In these cases, C_1 changed its voicing character entirely, e.g., *okozhatnak* [okosɦɔtnɔk] 'they can cause' (2a in Figure 8). Assimilation must have operated in a regressive manner here, but the beginning of voicing may have been anticipated during the production of C_2, resulting from a relatively early start of voicing for the following vowel.

The third major group of devoicing includes two subtypes of realisation. In these cases the realisation consists of one or two sounds as expected, but (both of) the sound(s) are/is partially voiced. 4 tokens (8.5% of all devoicing sites) were realised in this way. In 3 cases C_1 and C_2 differed in at least two distinctive features, the output therefore consisted of two different partially voiced consonant realisations, e.g., *több héten* [tøbɦeːtɛn] 'several weeks-superessive' (3d in Figure 8). In this subtype of realisations, the voicing of the adjacent sounds may have affected the voicing of the consonants analysed. The remaining one occurrence was realised in one long partially voiced sound (3a in Figure 8) affected perhaps by a physical constraint of devoicing arising due to aerodynamic and articulatory reasons (see, e.g., Shadle 1997; Stevens 1998): *kalauz szigorúan* [kɔlɔuzːigoruːɔn]. This phenomenon may appear only in clusters where the two phonemes differ only in terms of voicing, and after

regular regressive assimilation the expected form would be realised as one long sound.

The 44 voicing sites of the read speech samples were realised as follows. In the first major group C_1 was partially voiced in 13 cases (29.5%). These occurrences could be sorted into two subgroups: in 4 cases (30.8% of the 13 tokens) C_2 was unchanged in its voicing, e.g., *gyermeke**k** **b**ukfencezni* [ɟɛrmɛkɛg̊ bukfɛnt͡sɛzni]) (1a in Figure 8). In 9 cases (69.2%) C_2 changed its voicing character unexpectedly, e.g., *túlzá**sb**a* [tuːlzaːʒ̊pɔ] 'exaggeration-illative' (1b in Figure 8). In the first realisation subtype, the assimilation may have acted partially, in the second subtype several factors may have arisen. Either the assimilation operated in a regressive and progressive manner as well, but a physical constraint of devoicing intervened (and was balanced off later during the articulation), or the reader carried out the assimilation only later, due to some reading difficulty, for example.

In the second major group, where C_2 was partially voiced, 25 tokens (56.08% of the 44 cases) of all voicing sites of the readings were included. This major group consists of two subgroups. In 11 cases, C_1 changed its voicing character, e.g., *telje**s** **b**izonyossággal* [tɛjːɛʒ̊ bizoɲoʃːaːgːɔl]) (44.0% of the 25 cases) (2a in Figure 8). This means that in these cases the assimilation operated, but the physical/aerodynamic-articulatory constraint of devoicing might have acted as well. In the second subtype of this group, C_1 was deleted, e.g., *bolto**kb**an* [boltob̊ɔn] 'in shops' or *zöldségek, gyümölcsök* [zøltʃeːgɛ ɟ̊ymøltʃøk] (2b in Figure 8). This realisation type occurred 14 times (56.0% of the 25 cases). In these cases we cannot decide whether the assimilation operated or not. C_2 might be partially voiced again due to the physical constraint.

Again, the third major group includes two subtypes of realisations. In these 6 cases (13.6%) the realisations consist of one or two sounds as expected, but (both of) the sound(s) are/is partially voiced. In four cases both heterorganic consonants were realised partially voiced, e.g., *háztartá**sb**an* [haːstɔrtaːʒ̊b̊ɔn] 'in household' (3d in Figure 8). Two other occurrences consisted of one long partially voiced sound: e.g., *zavaroka**t, d**aganatos* [zɔvɔrokɔd̊ːɔgɔnɔtoʃ]) (3a in Figure 8). In both subtypes the physical constraint of devoicing may have acted. In cases where two heterorganic consonants appeared, a partially balancing articulation may have caused the partially voiced character of the second consonant.

In the spontaneous speech data 255 partially voiced clusters were found. Most of them (224 tokens, 87.8% of the cases) were expected

to be devoicing and the remaining 31 tokens (12.2%) were expected to result in voicing. 141 cases out of the 224 devoicing tokens were found in within-word clusters (62.9% of the tokens) and 83 cases (37.1%) were in between-word clusters. The voicing sites showed a reverse tendency. 10 out of the 31 cases (32.3%) were found in within-word clusters, while 21 cases (67.7%) in between-word clusters. The realisation types could be grouped into three major groups again, in terms of which sound was partially voiced.

In the case of the devoicing sites of the spontaneous speech samples, the distribution of the realisations was as follows. The first major group included 183 occurrences (81.7% of the 255 cases), where C_1 was partially voiced. The three subtypes were sorted on the basis of the realisation of C_2. C_2 was unchanged in voicing 164 times (89.6% of the 183 cases) (1a in Figure 8), e.g., *hogyha* [hoɟ̊hɔ] 'if' or *egy filmet* [eɟ̊filmɛt]) 'a film-acc'. In these cases assimilation may have partially applied. C_2 was deleted 17 times (9.3%) (1c in Figure 8), e.g., *biztos* [biz̥oʃ] 'sure'. In this subtype either the assimilation operated partially or an articulatory constraint caused the occurrence of the partially voiced consonant. C_2 changed its voicing twice (1.1% of the cases) (1b in Figure 8): *úgyhogy* [u:ɟ̊ɦoɟ]) 'so'. In these two cases the second consonant was /h/. This phoneme may be realised voiced in some contexts, but mostly between two vowels (Gósy 2005). In these cases, it was perhaps the voicing of the preceding phonemes and the following vowel that might have resulted in voiced realisation.

In cases where only C_2 was partially voiced (22 cases, 9.8% of the devoicing tokens) two minor groups were found. In the first subgroup, the voicing of C_1 changed (2a in Figure 8), e.g., *úgyhogy* [u:c̥ɦoɟ̊]. In the first subtype of this group the assimilation may have operated, while a partially progressive effect, and in the case of /h/ the formerly mentioned effect of the following vowel, might have acted as well again. This subgroup consisted of 11 tokens (50.0% of the 22 cases). 11 times (50.0%) C_1 was deleted (2b in Figure 8), e.g., *egy kicsit* [ɛgit̠ʃit] 'a little-acc'. In the second subtype the devoicing constraint is supposed to have operated.

The third major group included 19 cases (8.5% of the tokens) from the devoicing clusters. These realisations could be sorted into four subtypes. In three cases (15.8% of the tokens), one long partially voiced sound occurred with a differing manner or place of articulation from the two phonemes (3b in Figure 8), e.g., *nagyszerű* [nɔd͡z̥:ɛry] 'cool, great'. The second subgroup includes the realisations where one long sound oc-

curs (where the two neighbouring phonemes differ only in voicing) (3a in Figure 8), e.g., *tudtuk* [tud̥ːuk] 'we knew' or *meg katona* [mɛɡːɔtonɔ] 'and soldier'. This type of realisations appeared 6 times (31.6% of the 19 tokens). In these two subtypes the assimilation may have operated partially. The third subtype consisted of two partially voiced sounds (3d in Figure 8), e.g., *hogyha* [hoɟ̥ɦɔ]. Five tokens of the devoicing sites (26.3% of the cases) belonged to this subgroup. In four of these realisations the second consonant was /h/. In these clusters the above-mentioned concomitances are supposed to have acted. In one case, C_2 is a /t/, where perhaps a progressive assimilation and the physical devoicing and the compensatory effort may have acted. Five times (26.3% of the 19 tokens) both consonants were realised but they could not be analysed separately, e.g. *biztos* [biz̥t͡os]) (3c in Figure 8). In these cases, assimilation may have acted partially.

The 31 voicing clusters of the spontaneous speech data showed a somewhat different distribution. Only C_1 was partially voiced in 12 voicing sites (38.7% of the 31 cases). C_2 was unchanged in voicing only three times (25.0% of the cases) (1a in Figure 8), e.g., *piac, de* [piʲɔd̥z dɛ] 'market, but'; C_2 changed its voicing 9 times (75.0% of the cases) (1b in Figure 8), e.g., *és boldog* [eːʒ̥ boldog] 'and happy'. In the first subgroup the assimilation may have operated partially on C_1, while in the second physical devoicing might have appeared.

14 cases were found where only C_2 was partially voiced (45.2% of the voicing tokens of the spontaneous subcorpus). In the first subgroup, the voicing of C_1 changed, e.g., *képzett* [kebz̥ɛtː] 'educated' (2a in Figure 8). In this subgroup the assimilation may have operated and the second consonant may have become partially voiced due to aerodynamic and articulatory factors. This subgroup included 8 tokens (57.1% of the 14 cases). In four cases (28.6% of the 14 tokens) C_1 was deleted (2b in Figure 8), e.g., *filmekben* [filmɛb̥ɛn] 'in films'. In these cases we cannot tell if assimilation operated or not. The second consonant may have become partially voiced for aerodynamic reasons. C_1 stayed unchanged twice (14.3% of the 14 cases), e.g., *török, de* [tørøk d̥ɛ] 'Turkish, but' (2c in Figure 8). In these cases either progressive assimilation operated partially, or no assimilation occurred and the second consonant became partially voiced due to the aerodynamic factors referred to above.

In the third major group, where both consonants appeared partially voiced, five cases (16.1% of the 31 tokens) from the voicing clusters were found. Three tokens consisted of two different partially voiced sounds

(3d in Figure 8), e.g., *robot**b**ól* [robo d̪b̥o̥ːl] 'from hard work'. In two cases (where the two neighbouring phonemes differed only in voicing) a single long sound was realised: *szé**p** **b**őr* [seː b̥ːør] 'nice leather') (3a in Figure 8). In the first case the assimilation may have operated but the constraint and the compensatory effort may have appeared at the same time; while in the second subtype the assimilation may have applied totally but the constraint acted as well.

4. Discussion

4.1. CCC clusters

Overall, 128 CCC clusters were found in the material: 50 in reading aloud and 78 in spontaneous speech. The most frequent tokens were ones in which the third consonant was a sonorant or *v:* 50 in the read and 60 in the spontaneous subcorpus. In both subcorpora, the clusters were realised regularly most of the time (cf. Figure 7), i.e., voicing assimilation took place as expected and all three consonants appeared on the surface. Both subcorpora contained realisations in which one of the phonemes was deleted; however, the ratio of these forms differed. The most salient difference is that the cancellation of voicing assimilation (no VA) was attested only in spontaneous speech (however, the subgroup of C_1 deletion was larger in reading than in spontaneous speech).
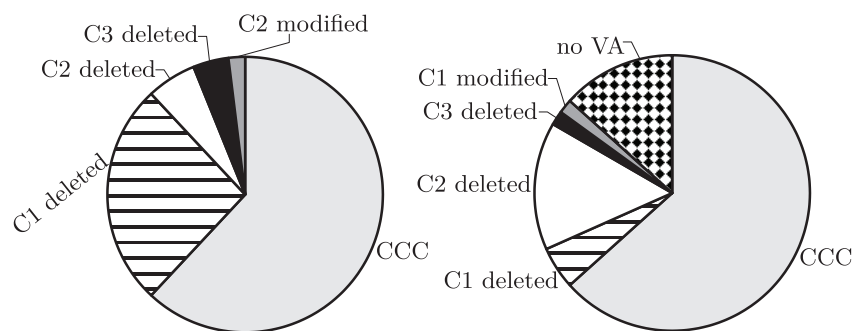

Fig. 7

The ratio of categories of CCC realisations (where the third consonant
is a sonorant or *v*) in reading aloud (left) and spontaneous speech (right)

Several cases were found in the read subcorpus where C_{obstr1} was deleted in a position preceded by a C_{son}. This type of occurrences is associated with casual and usually fast speech (cf. Siptár 1991), although our read material can be considered guarded speech on the basis of the subjects' intention. The articulation tempo of utterances in which these deletions occurred ranged between 11.9 and 14.5 sound/s with a mean of 13.3 sound/s, which corresponds to the Hungarian normal tempo (Gósy 2004). Therefore, these occurrences could not be caused by speaking fast. It should also be noted that in these cases the obligatory assimilation of nasals in terms of place of articulation did not operate between /n/ and /b/, but rather between /n/ and the disappeared velar plosive: *hazá**nkb**an* [hɔzaːŋbɔn] 'in our country', as opposed to a potential fast-speech version [hɔzaːmbɔn]—which, however, would be homophonous with *hazámban* 'in my country'. So the lenited form preserves the original feature of velarity in order that the perceptual system can recognise the structure of the word. Since reading aloud is traditionally considered a formal speech style, and the subjects were not trained speakers/readers, the question arises whether the standard norm of reading is getting closer to colloquial/casual speech. This can be the subject of future research.

4.2. CC clusters interrupted by pauses

The comparison of results found in read and spontaneous speech shows that the planning processes operating in these two mechanisms produce similarities as well as divergences. The relatively frequent appearance of hyperarticulation and especially of phrasing in the read material convincingly supports that in this speech style the phonetic planning and articulation processes are conscious to a large extent. Meanwhile—quite expectedly—planning difficulty is the most frequent motivation of the appearance of pauses between C_1 and C_2 in spontaneous speech. Phrasing operates here, too, but its ratio is relatively small in comparison to planning-difficulty-motivated cases. It is also remarkable that in spontaneous speech it is far more frequent that the voicing assimilation process operates in spite of a pause between C_1 and C_2. This fact shows that voicing assimilation is a fairly automatic process of speech planning.

Pause duration data in cases of cancelled assimilation show that phrasing is carried out similarly in reading aloud and in spontaneous speech. Hyperarticulation was supposed to be realised only in reading

and indeed, in comparison to the categories of spontaneous data, similarities cannot be observed. Likewise, the pausing parameters of planning difficulties that emerged on the surface considerably differ from the other groups' data. The last category of spontaneous material, where the reason of pausing and cancelled voicing assimilation could not be defined, is similar to the phrasing data. However, phrasing is associated with specific points of the syntactic-semantic structure, and these realisations do not fit into them (e.g., *tovább | folyt* 'continued', *lesz | gépekre* 'will be given for machines'). We still consider the two most probable reasons of these cases to be (i) non-surfaced planning difficulties or (ii) the speaker's individual articulation habits.

In the spontaneous speech material, the pauses of various durations between the words resulted in a lack of voicing assimilation in 82.8% of the cases. This is quite similar to Gósy's (1999) 77.2% result. In that study she found that pause duration influences the activation of phonological coding: if the pause is shorter than 55 ms, voicing assimilation always operates, but if the length exceeds 314 ms, it never takes place. Among our spontaneous data the application of voicing assimilation was associated with the range of pause duration between 105 and 767 ms. The huge difference between Gósy's cited data and the upper limit value found in our material confirms that the possibility of individual word-final devoicing should be taken into account. It is interesting, however, that in reading aloud voicing assimilation took place even in the case of a pause realised in 277 ms within a word. This occurrence suggests that, for example, a slip of the tongue can produce a relatively long pause while phonological encoding still results in the expected assimilation on the surface.

4.3. Partially voiced output of voicing assimilation

Among partially voiced occurrences 91 tokens (26.3% of the cases) appeared in the reading task and the remaining 255 (73.7%) in the spontaneous speech samples. Comparing the two speech styles the results show that (at least one) partially voiced consonant occurred at every 22nd (reading) or 25th (spontaneous speech) potential voicing assimilation site (consisting of two obstruents and interrupted by no pause). Since this difference seemed to be less marked than the effect of the expected type of the voicing assimilation, the results are to be discussed according to the latter distinction as well.
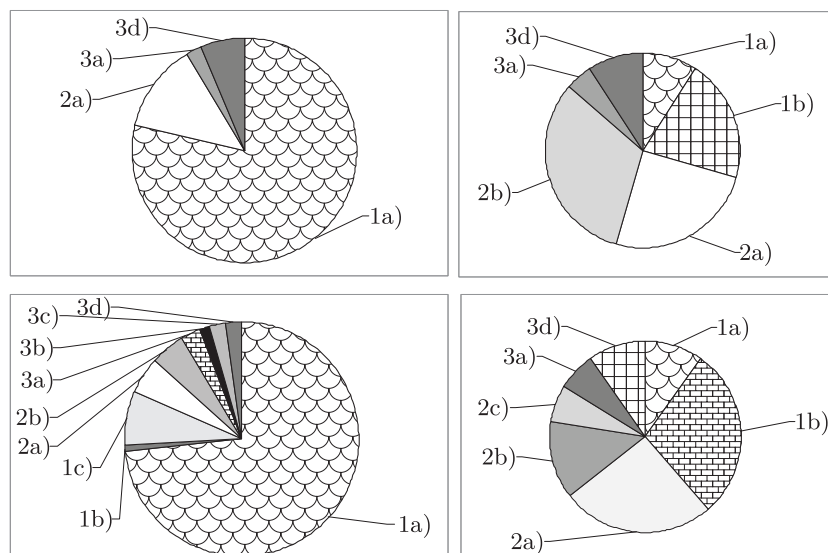
Fig. 8

The ratio of the subcategories of the partially voiced realisation type regarding speech style and type of assimilation (above: reading, below: spontaneous speech; left: devoicing, right: voicing).
(1a) C_1 partially voiced, C_2 unchanged, (1b) C_1 partially voiced, C_2 assimilated, (1c) C_1 partially voiced, C_2 deleted, (2a) C_1 assimilated, C_2 partially voiced, (2b) C_1 deleted, C_2 partially voiced, (2c) C_1 unchanged, C_2 partially voiced, (3a) long partially voiced C, (3b) C_3 partially voiced, (3c) $C_{1,2}$ partially voiced, (3d) both C_1 and C_2 partially voiced.

The three major groups of realisations showed a similar distribution for both assimilation types in reading and in spontaneous speech (Table 1). Three subtypes of realisations appeared only in spontaneous speech (Figure 8). These realisations appeared less than five times in all categories of assimilation and speech styles. The largest differences observed in the ratio of subtypes occurring in both speech styles are as follows. In the case of devoicing sites, the ratio of realisations where C_1 was assimilated regressively (as expected) but C_2 was partially voiced was higher than in the voicing sites. In both speech styles the ratio of these realisations was above 70% in the case of devoicing sites, while in the case of the voicing sites it stayed under 10% of all the clusters. C_1 deletion differed between the speech styles. In the read samples it was more frequent in the voicing clusters (above 30%, and no deletion was found in the devoicing clusters), while in the spontaneous speech samples the devoicing clusters showed

more C_1-deletion (12.9 vs. 4.9%). The subtype where C_1 was assimilated but C_2 was partially voiced was more frequent among the voicing clusters in both speech styles (reading: 25% vs. 12.8%, spontaneous: 25.8% vs. 4.9%)

Table 1

The ratio of the major groups of the realisation types in reading and spontaneous speech by both the devoicing and voicing types of assimilation

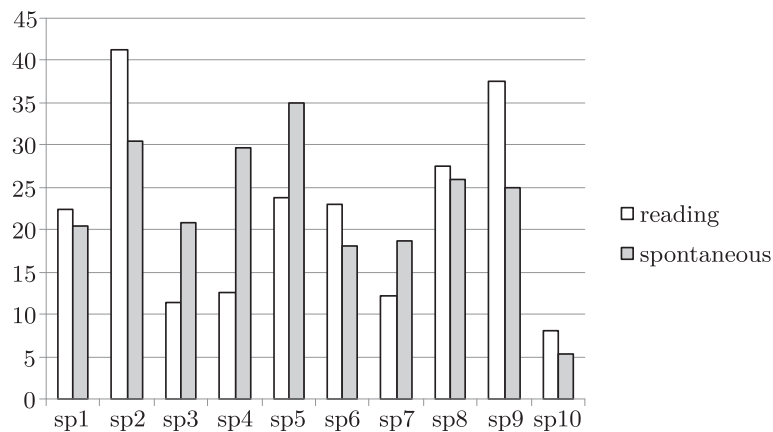|	Spontaneous		Reading	
	Devoicing	Voicing	Devoicing	Voicing
C_1 partially voiced	81.7	38.7	78.7	29.5
C_2 partially voiced	9.8	45.2	12.8	56.8
$C_1 C_2$ partially voiced	8.5	16.1	8.5	13.6

Fig. 9

Frequency of partially voiced realisations: inter- and intraspeaker differences

The frequency of partially voiced realisations showed relatively high inter- and intraspeaker differences as well (Figure 9). Regarding the speakers individually, the mean frequency was 22.0% in the read speech samples with an SD of 8.7%, while in the spontaneous samples this mean was 22.9% with 6.3% SD considering all possible voicing assimilation sites (where the cluster consisted of two obstruents and was not interrupted by any length of pause). Intraspeaker differences between the speech styles were also considerable. The highest difference was found for speaker 4:

only 12.5% of his read clusters were realised partially voiced, but in his spontaneous speech more than twice as many: 29.7% of the clusters included at least one partially voiced sound. In half of the speakers it was reading, in the other half it was spontaneous speech that more frequently contained partially voiced realisations.

70.0% of the partially voiced sounds (238 cases out of 340) showed a voiced part ratio shorter than 40% (but longer than 20%) of the total sound duration. Therefore we submitted the distribution of these cases to analysis (Figure 10). The ratio of these occurrences was expected to be smaller in the case of voicing sites as a result of an articulatory constraint. If the voiced part of C_1 is between 20 and 40% in voicing sites, this possibly means that the assimilation did not take place. In reading, the frequency of this kind of realisation is approximately 50%, whereas in spontaneous speech approximately 40% of all the voicing cases belonged to this category. On the contrary, if in cases of expected voicing the voiced part of C_2 is between 20 and 40%, this may mean that voicing assimilation acts in a progressive manner (as well). In reading, these cases reached approximately 20%; in spontaneous speech they exceeded 50% of all voicing cases. If the voiced part of C_1 is between 20 and 40% at devoicing sites, voicing assimilation must have applied. This happened in more than 70% of the cases in spontaneous speech, and in more than 50% in the read samples. If the voiced part of C_2 is between 20 and 40% at devoicing sites, progressive assimilation cannot be supposed, as we found in around 50% of the cases in both speech styles.
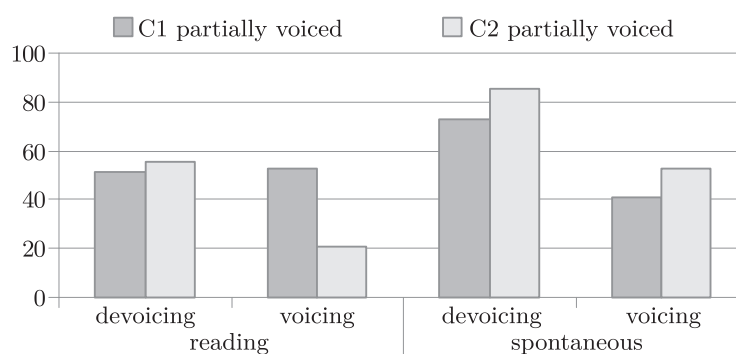

Fig. 10

Frequency of realisations with a voiced part between 20 and 40% regarding the type of assimilation, speech style, and the place of the consonant in the cluster

The results show that voicing and devoicing sites resulting in realisations including at least one partially voiced sound act differently in some respects, and speech styles differ in their realisations as well.

5. Concluding remarks

We suppose that the voicing assimilation dissimilarities of the two speech styles can be traced back to the differences of speech planning. During reading aloud, the speaker need not care about higher level planning processes. In speech production of this kind, macroplanning and some parts of microplanning are not involved (Levelt 1989). Since the subjects had time for reading over the text before reading it aloud, we can suppose that the production of the recorded speech was preceded by the interpretation of the text. Therefore, irregular outputs occurred in a higher ratio in spontaneous speech samples than in reading aloud. Phrasing overrode the operation of the phonological process in many cases in reading. On the other hand, some speech planning difficulty was the most frequent motivation of the appearance of pauses between C_1 and C_2 in spontaneous speech.

In some respects, considerable individual variation was observed. The assumption that word-final devoicing sporadically occurs in Hungarian speech can be supported by our data. The appearance of partially voiced realisations and its subtypes also varied across speakers. These results suggest that the voicing assimilation process can be an issue not only for phonological and phonetic research but also for psycho- and sociolinguistic studies and forensic phonetics.

In terms of partially voiced realisations, a large difference was found between the voicing and devoicing types of assimilation. This difference may result primarily from various articulatory concomitances like the interaction of voicing assimilation and the physical constraint of devoicing.

Regarding our starting assumption that spontaneous and read speech represent two different intervals of the H&H scale, the results suggest that reading aloud is really located closer to maximally carefully articulated hyper-speech than spontaneous speech is, see for example the relatively frequent appearance of hyperarticulation there. However, the standard norm of reading is probably getting closer to colloquial/casual speech, as the examples of fast cluster simplification showed. (Similarly, on the basis

of an analysis of Swedish TV-news, Torstensson (2004) concluded that reductions occur in scripted, carefully read newscast speech, too; however, the general impression of read speech is that it is close to the canonical forms.) Spontaneous speech variants nevertheless can be considered typical manifestations of hypo-speech—in various degrees.

The variability of the data confirms the scientific hypothesis that Hungarian voicing assimilation is a gradient and sometimes only partly regressive process. Even if it operates mainly obligatorily, several factors can override it. Hungarian voicing assimilation shows (among other things) some speech style dependent features.

References

Barry, Martin C. 1992. Palatalisation, assimilation, and gestural weakening in connected speech. In: Speech Communication 11:393–400.

Blaho, Sylvia 2008. The syntax of phonology. A radically substance-free approach. Doctoral dissertation, University of Tromsø.

Boersma, Paul – David Weenink 2005. Praat: Doing phonetics by computer. (Version 5.0) [Computer program]. http://www.praat.org/

Bóna, Judit – Tekla Etelka Gráczi – Alexandra Markó 2008. Coarticulation rules and speaking style dependency. In: Rudolph Sock – Susanne Fuchs – Yves Laprie (eds): International Seminar on Speech Production, 245–8. INRIA, Strasbourg.

Burton, Martha W. – Karen E. Robblee 1997. A phonetic analysis of voicing assimilation in Russian. In: Journal of Phonetics 25:97–114.

Gow, David W. Jr. – Aaron M. Im 2004. A cross-linguistic examination of assimilation context effects. In: Journal of Memory and Language 51:279–96.

Gósy, Mária 1999. Predictability of voicing assimilation in speech production. In: John J. Ohala – Yoko Hasegawa – Manjari Ohala – Daniel Granville – Ashlee C. Bailey (eds): Proceedings of the 14th International Congress of Phonetic Sciences (San Francisco), Vol. 3, 1905–8. University of California, Berkeley.

Gósy, Mária 2002. Temporal coding of voicing assimilation in speech production. In: Acta Linguistica Hungarica 49:257–76.

Gósy, Mária 2005. Pszicholingvisztika [Psycholinguistics]. Osiris Kiadó, Budapest.

Gósy, Mária 2008. Magyar spontánbeszéd-adatbázis – BEA [Hungarian spontaneous speech corpus – BEA]. In: Mária Gósy (ed.): Beszédkutatás 2008 [Speech research 2008], 194–207. MTA Nyelvtudományi Intézet, Kempelen Farkas Beszédkutató Laboratórium, Budapest.

Jansen, Wouter 2007. Phonological 'voicing', phonetic voicing and assimilation in English. In: Language Sciences 29:270–93.

Jansen, Wouter – Zoë Toft 2002. On sounds that like to be paired (after all): An acoustic investigation of Hungarian voicing assimilation. In: SOAS Working Papers in Linguistics 12:19–52.

Kenesei, István – Robert M. Vago – Anna Fenyvesi 1998. Hungarian. Routledge, London & New York.

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

Levelt, Willem. J. M. 1989. Speaking: From intention to articulation. MIT Press, Cambridge MA.

Lindblom, Björn 1990. Explaining phonetic variation: A sketch of the h&h theory. In: William J. Hardcastle – Alain Marchal (eds): Speech production and speech modeling, 403–40. Kluwer, Dordrecht.

Recasens, Daniel – Maria Dolores Pallarès 2001. Coarticulation, assimilation and blending in catalan consonant clusters. In: Journal of Phonetics 29 : 273–301.

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), 23–50. JATEPress, Szeged.

Shadle, Christine H. 1997. The aerodynamics of speech. In: William J. Hardcastle – John Laver (eds): The handbook of phonetic sciences, 33–64. Blackwell, Oxford & Malden MA.

Siptár, Péter 1991. Fast-speech processes in Hungarian. In: Mária Gósy (ed.): Temporal factors in speech, 27–61. HAS, Research Institute for Linguistics, Budapest.

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

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

Szigetvári, Péter 1998. Voice assimilation in Hungarian: The hitches. In: The Even Yearbook 3 : 223–36. http://www.btk.elte.hu/delg/publications/even/index.html.

Torstensson, Niklas 2004. Newstalk: Reductions in careful, scripted speech (Proceedings of Fonetik 2004, Department of Linguistics, Stockholm University). http://www.ling.su.se/fon/fonetik_2004/torstensson_fonetik2004.pdf

Vago, Robert M. 1980. The sound pattern of Hungarian. Georgetown University Press, Washington DC.