Coarticulation Rules and Speaking Style Dependency

Judit Bóna*, Tekla Etelka Gráczi** and Alexandra Markó*
*Eötvös Loránd University, **Research Institute for Linguistics, HAS
E-mail : bona_judit@freemail.hu, graczi@nytud.hu, markoxa@ludens.elte.hu

Abstract

This study examines voicing assimilation in different speaking styles in Hungarian: reading, interpreted and spontaneous speech. CC phoneme clusters consisting of voiced and unvoiced obstruents (in any order) were selected and were analyzed in terms of the presence of voice. It was investigated whether the operation of regular voicing assimilation is affected by the speaking styles, the rate of articulation or reading proficiency (based on the fluency of the reading and errors). The results show that speaking styles affect regular voicing assimilation less, while speaker variability is remarkable in the process. Among the individual characteristics, the rate of articulation and reading proficiency do not influence coarticulation at all. According to the results regular voicing manifests at the same rate as de-voicing. Results suggest that Hungarian voicing assimilation rules operate in praxis between 75–100% of the cases. Considerable individual variability makes the results suitable for use in forensic phonetics. In addition, better and more precise understanding of coarticulation can enhance the development of more natural-sounding synthesized speech.

1. Introduction

Hungarian has a phonological rule of voicing assimilation (VA) whereby obstruent clusters come to share the voiced/voiceless specification of their rightmost member [1]. Voicing assimilation is postlexical, it applies across word boundaries as well as within words and has two types. (i) The first one occurs when the voiced consonant follows a voiceless one and (ii) the second one when the voiceless consonant follows a voiced one. Voicing assimilation's direction is always backward.

Therefore VA in Hungarian speech is often considered as a purely regressive and categorical phenomenon. However, in practice divergent realizations can be observed [2, 3].

The majority of the earlier investigations were carried out on read sentences. The VA in Hungarian spontaneous speech was examined by Gósy [4]. Her study focused on the operation of the rule at word boundaries in the case of pauses (in connection with the activation processes of the mental lexicon).

Our present study examines VA on CC clusters in three different speaking styles: reading, interpreted and spontaneous speech. These speaking styles can be characterized by differences in the speech planning. Based on the experiences of former research on Hungarian spontaneous and read speech, our hypotheses are the following. Individual variation between speakers is greater than variation between speaking styles. In terms of VA realizations, spontaneous speech contains the largest proportion of irregularity, and in this speaking style the operating of VA rule depends on the rate of articulation (as some authors suggest, e.g. [5]). Whilst reading, the script affects the realization of VA due to the fact that Hungarian spelling is quasi phonologic, the letters can correspond to phonemes. Therefore reading proficiency influences the VA results. In general, voicing is less operative as a rule than devoicing.

2. Method

Ten Hungarian adult native speakers (5 female and 5 male) between the ages 29 and 64 participated in the research. Three different types of speech samples were recorded from each participant in a studio.

One task was reading aloud a 13 sentences long newspaper text and 14 independent sentences. In this material the number of the potential VA places was 48, but in reading aloud some of them were omitted by the speaker (because of misreading/uncertainty or prosody). Consequently 37-44 C_1C_2 clusters were realized without pause by the different speakers.

The other speech sample type was quasi-spontaneous or interpreted speech in which the participants had to convey the contents of a recorded text they listened to. The length of these samples were between 1.5 and 4 minutes, this subcorpus contains approximately 30 minutes speech material. In

the different speakers' productions the numbers of the VA places were between 10 and 34.

The third recording was that of spontaneous speech from discussions on several topics. This subcorpus contains texts between the length of 2.5-14 minutes (almost one hour altogether), where potential VA places occurred between 32 and 127 per subject.

All CC phoneme clusters of both within word (C_1C_2) and across word boundaries ($C_1\#C_2$) which were not interrupted by pauses of any length and consisted of voiced and unvoiced obstruents (in any order) were selected from the recorded material. The number of analyzed C_1C_2 clusters was 1190 in the total corpus (from ten speakers, in three speaking styles, cf. Table 1). This number includes 68 types of C_1C_2 clusters. 70.6% of the tokens represents voicing process, while devoicing occurs only in 29.4% in the corpus.

Table 1. *Distribution of C_1C_2 clusters in the subcorpora.*

Subcorpus	Number of possible VA places	Ratio in the corpus (%)
Reading	397	33.5
Interpreting	216	18.1
Spontaneous	577	48.4

The voicing of the consonants appeared in graduation from total voicing up to total devoicing. Therefore the consonants were categorized as follows. Voicing character was determined on the basis of subjective and objective evaluation of both oscillogram and spectrogram data. The pulses detected by the Praat and the presence of the voicing part on the spectrogram were considered. A consonant was considered as voiced if it contained quasi-periodic signal in at least 80% of its duration. The consonant was considered as unvoiced if it contained quasi-periodic signal in at most 20% of its duration. Between these values the consonant was evaluated as semi-voiced.

The text of the prosody units (marked from pause to pause) in which the C_1C_2 clusters realized, was annotated, and the rate of articulation was defined in sound/second.

Reading proficiency was subjectively evaluated on the basis of fluency, misreadings, interpretation and prosody. Based on these criteria, marks were given between 1 and 5, where 1 means the worst and 5 the best performance.

Three speaking styles (reading, interpreting and spontaneous) were compared with respect to the realization of the VA process. The ratio of regularly vs. irregularly operated assimilation was determined according to speaking styles and in terms of speakers as well. The modes of realizations were categorized. The relation between articulation tempo and the type of VA process operating as well as between reading proficiency and VA was examined.

For the acoustic analysis Praat 5.0 was used and the statistic analysis (one-way ANOVA) was carried out by SPSS 13.0.

3. Results

The consonant clusters of the total corpus were grouped into two main categories: (A) where the VA operated regularly, and (B) where the VA operated irregularly. The assimilation was considered regular if C_1 properly (at least in 80% of its duration) adapted to C_2's voiced/voiceless specification. Otherwise the cluster was categorized into group (B). The ratio of (A) and (B) types in the corpus was 88.5% and 11.5%, respectively.

The realizations can be ordered into seven subcategories (for the examples see Table 2). In type (A) three possibilities were found:

(i) The VA was regular (C_1 assimilated) and resulted in the expected sounds in 85.7% of the data.

(ii) The VA was regular (C_1 assimilated), but C_2 was deleted in 1.2% of the corpus.

(iii) In 1.6% of the cases the assimilation operated (the result is voiced or unvoiced according to the rules), but the output is different from the expectations regarding the manner and/or place of articulation of the target sound.

In type (B) four variants appeared:

(iv) The assimilation process did not operate at all in 3.1% of the corpus.

(v) The VA was progressive in 1.3% of the cases.

(vi) The result is semi-voiced in 4.5% of the data. In some of the cases, C_1 was partially voiced, together with C_2 as well, therefore the VA is irregular.

(vii) C_1 deletion occurred in 2.6% of the cases.

Table 2. *Examples of VA realization types.*

Category	Phonological form	Expected phonetic form	Realized phonetic form
(i) Norm. VA	biztoʃ	bistoʃ	bistoʃ
(ii) C_1	hoɟhɔ	hochɔ	hocɔ
(iii) C_3	ɔzt	ɔst	ɔsts
(iv) No VA	mɛgtɔrtɔni	mɛktɔrtɔni	mɛgtɔrtɔni
(v) Progr. VA	keːpzeːʃ	keːbzeːʃ	keːpseːʃ
(vi) Semivoic.	uːɟhoɟ	uːchoɟ	uː̥ɟhoɟ
(vii) C_2	ɛɟ kitʃit	ɛc kitʃit	ɛ kitʃit

The ratio of voicing and devoicing showed no difference. Voicing takes place regularly in 85.8%, and devoicing in 89.3% of the tokens. Therefore in the following the results of the VA-realizations are discussed without differentiating for these two kinds of the VA.

Figure 1 shows the distribution of the VA realization categories in the corpus.

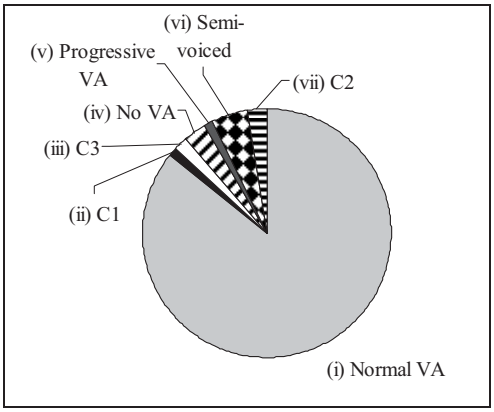

Figure 1: *VA subtypes in the corpus*

In terms of types (A) and (B) we did not find remarkable differences in the mean of the results between speaking styles. Nevertheless, Figure 2 shows that the distribution of the regularly realized VA-s are different.

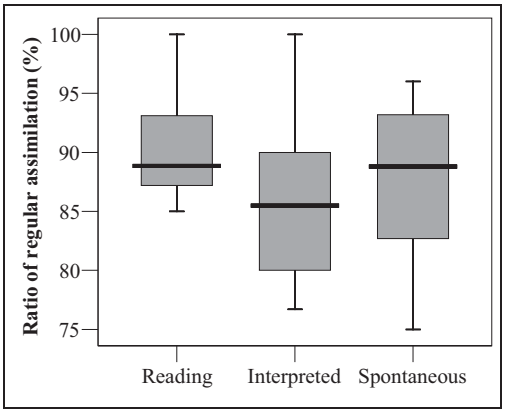

Figure 2: *Ratio of regular VA in the different speaking styles*

The distribution of categories (i)-(vii) in the different speaking styles is rather diverse (Figure 3).

The group (vi) was equally frequent in the examined speaking styles. On the other hand, C_2 deletion (ii) and the so called unexpected output (iii) predominantly occurred in the spontaneous subcorpus. This phenomenon can be explained by the informal characteristics of this speaking style.

The differences among the speakers (Figure 4) do not signal any tendency, nor the ages, neither the gender of the speakers influence the VA process.

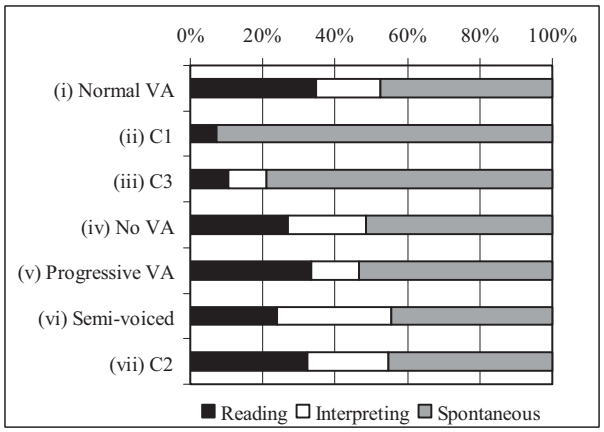

Figure 3: *Distribution of VA categories in the different speaking styles*

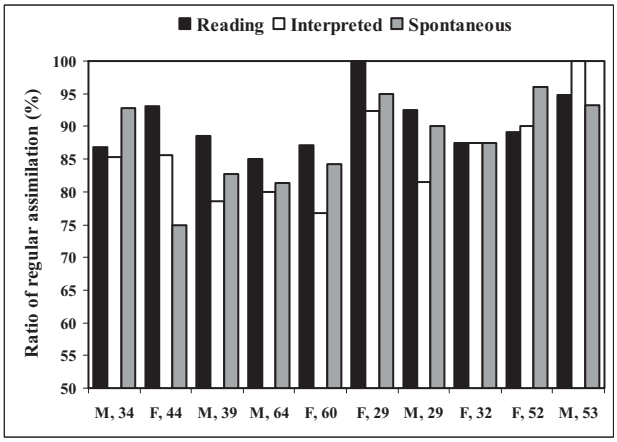

Figure 4: *Ratio of regular VA in the different speaking styles according to the speakers (M = male, F = female, numbers indicate the age)*

Investigation of the articulation rate of spontaneous speech utterances showed that the subtype (ii) differs from the other categories (cf. Figure 5). In the examples of (ii) C_2 deletion can be caused by the relatively fast tempo of articulation, which was confirmed by the post hoc test. While the one-way ANOVA's result was not significant [F (6, 570) = 1.993, p = 0.065), according to the Tukey test subgroup (ii) significantly differs from subgroup (i) and subgroup (iv) (p is 0.037 and 0.033, respectively).

It can also be seen that even in the case of extremely fast or slow articulation rate, the regular and expected VA is the most probable realization, which means that the tempo does not influence VA process to a great extent.

We compared the evaluation of reading level and the frequency of regular assimilation in reading according to the subjects (Table 3). In contrast to our hypothesis, no correlation was detected. Although the participant ("F, 29") whose reading was evaluated with the highest mark (5), realized 100% of the potential VA places according to the phonological

rules, some other speakers – even if they realized the C_1C_2 clusters regularly in relatively large ratio – were evaluated much lower (see for example "F, 44" or "M, 53"). This result means that the VA process is not influenced by the reading skill on the one hand, and, on the other hand, our subjective impression about someone's reading proficiency not necessarily depends on the quality of C_1C_2 realizations.

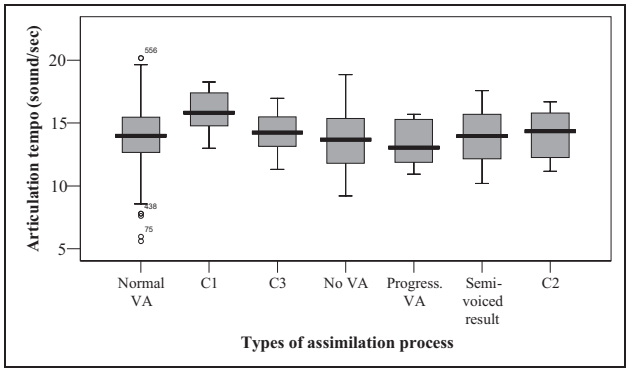

Figure 5: *The relation between articulation tempo and VA realization categories*

Table 3. *The relation of reading proficiency and VA realizations in reading.*

Speaker identifier (gender, age)	Reading proficiency (1 to 5)	Ratio of regular VA in reading (%)
M, 34	3	86.9
F, 44	1.5	93.1
M, 39	3	88.5
M, 64	4	85.0
F, 60	4	87.2
F, 29	5	100.0
M, 29	4	92.5
F, 32	4	87.5
F, 52	1.5	89.2
M, 53	3	94.8

5. Conclusions

The aim of the present research was to define the characteristics of the VA process in different speaking styles. For spontaneous speech similar data were not available from the previous studies, therefore in this respect our corpus's nearly 600 spontaneous data are useful also for the further research.

Although variation between speaking styles has not been proved by the means of the regular VA-realizations, the hypothesis that the highest proportion of irregularity can be experienced in spontaneous speech has been verified on the basis of the distribution of the data. The reasons behind this phenomenon are the characteristics of speech planning. The examples of category (iv) (where VA does not operate) confirm that irregular VA realizations can be explained by the speaker's uncertainty in the continuation. In this case a further methodological problem arises: how to categorize those realizations, which can be traced back to a disfluency phenomenon.

The great differences of some categories' appearance among the subcorpora suggest that examination of VA can add information to sociolinguistic studies in terms of formal-informal styles/registers.

Individual variation between speakers is remarkable, namely semi-voiced realizations seem to be individual articulation properties. Nevertheless the operating of VA rules in spontaneous speech does not depend on the rate of articulation, and reading proficiency does not influence the operating of VA rules in reading. The data from the reading subcorpus confirm that this postlexical phonological rule is automatic to a great extent, so the quasi phonologic Hungarian spelling does not "overwrite" it even if the reader's proficiency is relatively poor.

According to the results regular voicing manifests at the same rate as de-voicing. Results suggest that Hungarian voicing assimilation rules operate in 75–100% of the cases. Further studies have to give an explanation for every single case in which the VA process is irregular. The results suggest that to give answer to the questions about VA process is an issue not only for phonological-phonetic research, but also for psycho- and sociolinguistic studies. Considerable individual variability makes the results suitable for use in forensic phonetics. In addition, better and more precise understanding of coarticulation can enhance the development of more natural-sounding synthesized speech.

References

[1] P. Siptár, M. Törkenczy, *Hungarian phonology.* Cambridge University Press. Cambridge, 2000.

[2] D. W. Gow, A. M. Im. A cross-linguistic examination of assimilation context effects. *Journal of Memory and Language,* 51:279-296. 2004.

[3] W. Jansen, Z. Toft. On sounds that like to be paired (after all): an acoustic investigation of Hungarian voicing assimilation. *SOAS Working Papers in Linguistics,* 12: 19-52. 2002.

[4] M. Gósy. Temporal coding of voicing assimilation in speech production. *Acta Linguistica Hungarica* 49(3-4): 257-276, 2002.

[5] R. Vago, *The Sound Pattern of Hungarian.* Georgetown University Press, Washington, DC, 1980.