

ON THE REDUCTION OF CONSONANT CLUSTERS IN POLISH ACCORDING TO EMOTIONAL SITUATION

P. Durand* and R. Gubrynowicz†

* *Laboratoire "Parole et Langage", Université de Provence, Aix-en-Provence, France,*

† *Speech Acoustics Laboratory, Institute of Fundamental Research of Technology, Warsaw, Poland*

ABSTRACT

The present study investigates the effect of speaking style on the temporal structure of fricative-affricate clusters. It is well known that duration of vocalic as well as non-vocalic portions of speech are affected by a large number of linguistic and non-linguistic factors (e.g. stress and speaking rate) To study the reduction phenomena in fricative-affricate clusters the Polish BABEL Database [2] was applied. The main part of this database is composed of 40 short passages related to different "normal" speech situations, such as telephone calling, narration or inquiry. Recordings of four speakers with different characteristic speaking rates were analyzed and described. The preliminary results concerning the importance of such factors affecting fricative-affricate cluster timing as the rate of speaking and stress are given .

1. INTRODUCTION

The present study is a part of investigations related to analysis of factors affecting articulatory timing and reduction in Polish consonant sequences. Polish is well known for its rich consonantal system, mainly for its opposition between simple and palatalized articulation in fricatives and affricates, both for their unvoiced as well voiced versions. In Polish consonantal system there are 5 unvoiced fricatives, 4 voiced fricatives and 3 affricates in unvoiced and voiced versions (Fig. 1). As for an inventory of 37 Polish phonemes, the combined frequency of occurrence of fricatives and affricates is relatively high (19,34% - and 12.2% for unvoiced version, after the data in [1,2]).

	Labio-dental	Alveolar	Post-alveolar	Pre-palatal	Velar
UNVOICED FRICATIVES	f	s	ʃ	ç	x
VOICED FRICATIVES	v	z	ʒ	ʒ̣	
UNVOICED AFRICATES		t̪s	t̪ʃ	t̪ç	
VOICED AFRICATES		d̪z	d̪ʒ	d̪ʒ̣	

Figure.1 The Polish fricatives and affricates chart

Current spoken Polish is a language that admits very complex consonantal clusters both in syllable onset and rhyme [3]. Thus it is not unusual to find very long sequences of consonants

whose first part is the coda of the preceding, and second, the onset of the following syllable. In word on initial and medial positions, consonantal clusters are either fully voiced, or fully voiceless (originally, or due to voicing assimilation). The clusters of fricatives and affricates are of special interest as they are frequently present not only on the word boundaries, but also inside the words. The most frequent sequence fricative-affricate is /Xɹ/ which within 1197 admissible Polish dyads [1] is after the frequency of occurrence on the 110 position. The cluster /Σʃ/ is on the 185 position, /σʃ/ - 350, /ξʃ/ -431, /ξç/ -445.

In this preliminary study, we intend to show some effects of speaking style and especially of tempo on acoustic realizations of Polish spoken dyads composed of fricatives and affricates. We want to give an overview of phenomena that take place when speech tempo increases due to changing of speaking style. Besides this, co-articulatory factors were investigated:

- the time reduction in consonantal groups composed of fricatives followed by an affricate consonant, with the same place of articulation,
- the time reduction in consonantal groups composed of fricatives followed by an affricate consonant, with a different place of articulation.

In our study, duration was applied as the main acoustic parameter characterizing the segmental structure of investigated consonantal clusters. To enable a comparison of realizations across speaking styles the sequences with a complete segment reduction were not taken into account.

2. MATERIAL AND METHODS

For this study we have selected recordings from Polish Babel Database [3] The selected speech material is composed of 40 short passages corresponding to various speech situations, such telephone conversation, short stories, inquiry etc. The mean length of passages is 56 words. Before starting to record passages, speakers were first instructed to acquaint themselves with each item separately and to adopt their style of speaking to the situation related to the text. However, it should be noted that speakers were strongly requested to not overdo it and to preserve their personal style of expression. The recorded passages of various emotional meaning could be divided into following groups: narration, a talk to other person, call by telephone, customer complaint, inquiry, monologue and reading report.

The influence of speech tempo on consonant cluster timing was investigated on the basis of the recordings made by four speakers (2 male and 2 female) of different characteristic speaking rate, ranging from 140 words/min to 208 words/min.

From the phonetic transcriptions of the prompting texts all fricative-affricate and affricate-fricative clusters were checked and located in speech recordings, also those spoken across the word boundaries. In total, 720 clusters were identified.

Phoneme boundaries were placed using a signal editor Phonedit 3.68 and analyzing the signal waveform, spectrum and energy. Though the boundaries of fricatives and affricates spoken in vocalic context are not difficult to be determined and the changes in the spectral composition of the waveform are there very distinct, for clusters, however, especially for a sequence affricate-fricative they could be in most cases placed only in an approximate way. But also for a sequence fricative-affricate, in principle easier to be segmented, in some cases it is not always sure to place the beginning and the end of the closure separating the two fricative segments. The origin of this difficulty is the incomplete closure of the vocal tract during the transition from the fricative to the following affricate. An example of such cluster realization is given on Fig. 2. It represents the cluster /Σʃ/ spoken in intervocalic position /Eɔ - I/, in the word "m• czyzna" - /μEɔΣʃIvα/ (a man). In this case the beginning and the end of the incomplete closure could be evaluated after spectral and envelope level changes.

However, there are some cases when the boundaries of the closure are not so evident (especially for their voiced counterparts) and for this reason, in the preliminary approach, the cluster duration was taken as a whole for the analysis, and the inside timing was analyzed only when the boundaries were unquestionable. Particularly difficult clusters to segmentation are sequences affricate-fricative where in most cases it is impossible to distinguish the fricative segment belonging to the affricate from this one of the following fricative consonant.

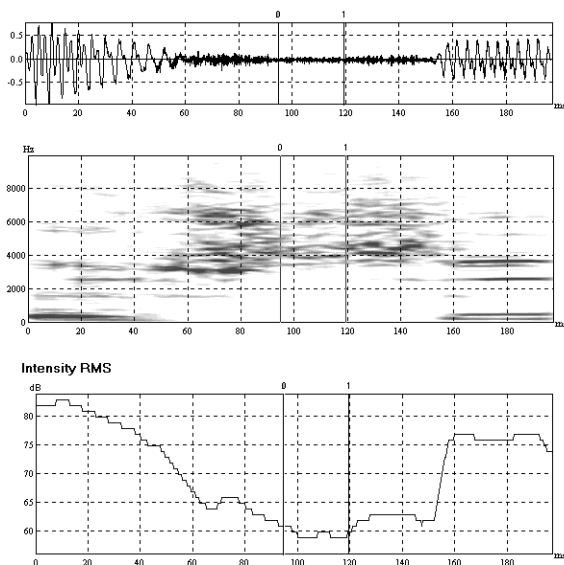


Figure. 2 An example of speech waveform, spectrogram and level envelope of a cluster /Σʃ/ spoken with incomplete closure. The boundaries of the closure are marked with vertical continuous lines (0,1).

3. RESULTS

The analysis of the speaking rate of the recorded passages shows that when speaker tries to speak with more emotion he/she generally speaks faster, but at the same time, very often, more focal stresses occurs in the utterance. They involved, in general, the raise of the stressed syllable duration. From our data we observed that for some texts the rate of speaking is higher 15-20% than the mean speaker's specific rate, for other texts like reports, narration without emotional climate, the rate of speaking observed in our data is lower, of about 10-16%.

From the analysis of data across speakers it is evident that duration of fricative-affricate sequences are strongly related to the speaker's characteristic rate of speaking. Although this correlation is not linear, it is clear that when the speaker's characteristic rate of speaking is higher, the duration of the cluster is in most cases shorter. A similar change is observed for every speaker when he varies his rate of speaking. However, in analysis of duration within intra-speaker data the influence of speaking rate is not always so distinctive. This is due to other factors affecting duration, like stress or cluster position.

Polish is a language with word stress placed on the penultimate syllable. There are, however, some categories of departures from this general rule. This question is more developed in [5]. As to sentence prominence, similarly to what happens in most languages, the decision of what part of the sentence to emphasize is motivated by the speaker's focus of interest or by his/her desire to contrast part of the sentence with a potential alternative. The difference between English and Polish is that, Polish being a free word order language, it often employs shifting the word to be stressed to the front of the sentence. Thus, in addition to the means available to speakers of English, Poles can use syntactic operations to express the focus of interest. This was taken into account in composition of texts of the database.

An important factor affecting the timing of fricative-affricate cluster is its position in the utterance. For example, when it is spoken on the beginning of the phrase the fricative segment duration is not so strongly related to speaking rate, although it varies across speakers but not in regular manner. It happens, for example, for the speaker of the fastest tempo of articulation the duration of the initial fricative is largely longer than for other speakers. This is observed, especially, when the place of articulation of the following affricate is very different from this one of the fricative, as it is in velar-alveolar cluster /ʃʒ/ The influence of the cluster position on the timing of the fricative-affricate cluster is strongly marked when this sequence is uttered on the end of the phrase. Both unvoiced fricative segments are longer, but frequently the affricate is so weak that it often undergoes total reduction. A tendency to increase duration of both fricative segments of the cluster occurs when it is spanning boundaries of the bordering words and the place of articulation is different. However, the effect of the change of the place of articulation within clusters uttered in inter-vocalic position seems to be relatively small.

Stress is another important non-linguistic factor affecting the fricative-affricate timing. The neighborhood of a stressed vowel can contribute in the duration increase of the whole cluster. This can frequently counterbalance the effect of the speaking rate raise. The influence of stress was analyzed, however, for a limited number of clusters, mainly for those composed of unvoiced fricative-affricate sequences. When a

stressed vowel precedes the cluster, it is mainly the first fricative segment which is affected by and its duration rise to well over double of the length observed for unstressed version. In case a stressed vowel follows the cluster its influence is not so evident. It seems that the pronunciation of the affricate is relatively stable and is not so dependent on speaking rate and stress. However, it must be emphasized that all these results are related to sequences of unvoiced fricative and affricate.

The clusters affricate-fricative are difficult to examine. The main problem is a division of two fricative segments not separated by an occlusion. A frequent assimilation or total reduction make this task very difficult. As the segmentation is in this case very uncertain these clusters (unvoiced as well in voiced version) were not included into investigations.

The voiced fricative-affricate sequences are much more difficult to be analyzed, because their articulation is frequently, unstable. The analysis of voiced clusters shows that even they are spoken by a speaker in an exactly same context (within the same word) their temporal structure is variable from utterance to utterance and is not very correlated with the speaking rate. In comparison to their unvoiced counterparts, it seems that the temporal structure of voiced clusters is more frequently affected by a total reduction of a one of the component segments. Often reduction is connected with assimilation, as it takes place, for example, in a sequence /ʒʁ/ which is uttered /ʁ/ or /ʒZ/ or simply /z/. These phenomena are of course more frequent in fast speech, but not only, and they are correlated to speaker's voice characteristics and strategies in speaking communication.

4. CONCLUSION

In this preliminary study we have presented that duration of fricative-affricate clusters are influenced by a large number of linguistic and non-linguistic factors (e.g., speaking rate and stress). The influence of the style of speaking was expressed mainly by examination of the effect of speaking rate on fricative-affricate cluster duration. However, there are many other factors connected to the style of speaking (stress timing, focal stress and others) which have influence on temporal structure of the utterance. But this phenomena are for Polish completely unknown. The results of our investigations demonstrate that cluster duration is related to speaking rate, stress, position, but it is difficult to describe these correlation in a more formal manner.

ACKNOWLEDGEMENT

The investigation was done in part, within the framework of the French-Polish common research project 3676 "PAROTEXT".

REFERENCES

- [1] Jassem W., Lobacz P., Phonotactic Analyses of Spoken Polish Text (in Polish), *Biul. Polskiego Tow. Językoznawczego*, z. XXXII, 1974
- [2] Rocławski B. *Phonotactics of the Contemporary Polish* (in Polish), Ossolineum, Warsaw, 1981
- [3] Gonet W., Duration of R.P. Monophthongs in monosyllabic words, *Speech and Language Technology*, (ed. Jassem W., Basztura Cz.), Polish Phonetics Ass., Poznan, 1997, 195-228.
- [4] Gubrynowicz R., The Polish database of spoken language, *Proc. 1st Conf. on Languages Resources and Evaluation*, Granada, May 1998, pp.1031-1037
- [5] Wierzbowska B., *Phonetics and phonology of Polish language* (in Polish), Ossolineum, Wrocław, 1980