ISCA Archive
http://www.isca-speech.org/archive

6th European Conference on
Speech Communication and Technology
(EUROSPEECH'99)
Budapest, Hungary, September 5-9, 1999

ACOUSTIC CHARACTERISTICS OF PLOSIVES IN CONSONANT-CONSONANT SEQUENCES AT WORD BOUNDARIES

Natalia Petlyuchenko
Odessa State University
Romance-Germanic Faculty
24/26 Frantsuzsky Blvd., Odessa, 270058, Ukraine
sesame@odessa.net

10.21437/Eurospeech.1999-112

ABSTRACT

In this paper we describe acoustic characteristics of aspiration, explosiveness and tension of plosives within the framework of consonant sequences at boundaries of two words extracted from German read speech. In these sequences (66 types) the prejuncture consonants (in final position in the preceding word) were limited to the plosives p/t/k, whereas the post-juncture consonants (in initial position in the following word) could consist of the entire set of German consonants. The proposed order of sounds at word boundaries enables to trace modifications of the plosives in question and degree of their correlation with a number of positional-combinatory and prosodic factors, namely type of following consonant, presence of pause inside the sequence, position of plosives in rhythmic groups with regards to tone syllable.

INTRODUCTION

Investigation of plosives was carried out within the framework of consonant sequences, located at word boundaries. A sound sequence represents a minimum unit of movement of speech organs and is formed in the process of continuos articulation, characterized by successive overlap of articulatory gestures. This causes certain difficulties in determining the structure of the acoustic signal of sound sequences, all the more, the position of sequences at word boundaries is connected with slowdown of articulation and appearance of pauses [3; 5; 11].

Voiceless final plosives in continuos speech are characterized by significant loss of aspiration, reduction of the explosive phase and voicing [1; 5; 7]. It is important to determine, whether these modifications are a result of general laxation of articulation, or they are connected with a number of positional and prosodic factors, which should be taken into consideration when describing the mechanism of formation of sound units in speech flux.

EXPERIMENT

The experimental material consisted of 2.5 hours of authentic German news texts spoken by eight (4 male and 4 female) German public radio and television newscasters.

About 1400 consonant-consonant pairs were extracted from these continuos speech samples and classified according to type of juncture. The prejuncture consonants (in final position in the preceding word) were limited to the plosives p/t/k, whereas the post-juncture consonants (in initial position in the following word) could consist of the entire set of German consonants: (1) homorganic plosives (2) non-homorganic plosives; (3) fricatives, affricates; (4) sonorants. We also considered the absolute ending position in sentence, i.e. when plosives p/t/k were not followed by any sound, or when the duration of the following pause was maximum (up to 1000 ms). This gave us a total of 66 types of sound sequences.

For segmentation of the plosives from the preceding vowels on the waveforms, the zero-crossing point after the last glottal pulse of the preceding vowel was taken as the beginning of the plosive. On the spectrogram the vowel was identified by increase (decrease) of spectral energy in the 1000 Hz range. Parameters of frequency and energy were measured at different sampling points in the highlighted segments.

As the studied consonant sequences were located at word boundaries, it was necessary to take account of possible slowdown of articulation and appearance of pauses. Interruption in speech can take place both within the sound limits (intrasegmental pause), and outside its limits (intersegmental pause). It the present research we adopted the classification of intersegmental pauses into *minimum pauses* (5-20 ms), *medium pauses* (20-60 ms) and *maximum pauses* (60-100 ms). On the perceptual plane, pauses were aurally divided in the same order into weakly distinguishable, normally distinguishable and highly distinguishable.

RESULTS AND DISCUSSION

Aspiration of plosives was described in temporal characteristics proceeding from the duration of this phase. It was observed that prejuncture plosives in final position are characterized by significant loss of aspiration (70%). Duration of intersegmental pause of more than 60 ms is significant for describing the degree of aspiration of plosives ($p < 0.01$). Here we traced clear correlation between the duration of the aspiration phase and duration of intersegmental pause ($r = 0.79$; $p < 0.001$). At this, the following consonant is insignificant ($p > 0.07$). Duration of pause of 5-20 ms leads to a reduction of the amount of aspirated plosives (28 %) or reduction of the duration of the aspirational phase to 10-15 ms. In 58 per cent of the cases, duration of pause at word boundaries of more than 40-60 ms can serve as a criterion of aspiration of plosives. When the duration of pause amounts to 60-100 ms, the duration of aspiration phase reaches 80-140 ms, and here we can count on a high percentage of aspiration (92 %). In the absolute ending position, when the plosive is not followed by any sound, and the duration of pause is maximum, the number of aspirated plosives amounts to 93.5 %. The number of junctures without intersegmental pause, but where aspiration is nevertheless present amounts to only 8 % of the cases, this indicating their sporadic character.

We also observed that aspiration does not depend on the position of plosive in final tone syllable ($p > 0.1$). In order to exclude the influence of pause which is a very strong factor, we compared sound sequence in which pause is absent. Moderate aspiration of plosives (40-50 ms) was present in both stressed and unstressed syllables.

Plosives in consonant sequences at rhythmic group boundaries were characterized by longer duration of the aspiration phase compared to junctures inside rhythmic groups ($p < 0.05$). Maximum duration of aspirational phase of prejuncture plosives at rhythmic group boundaries (55-100 ms) accounted for 52 % of the total amount rhythmic group boundaries, whereas such duration of aspirational phase in plosives located at word boundaries inside the rhythmic group accounted for only 15% of the cases.

Explosiveness of plosives was described in terms of duration and energy, proceeding from the duration of the explosive phase and maximum intensity in areas of burst energy concentration. It was observed that duration of explosive phase in plosives amounts to an average of 10-12 ms, this being a mean value for all sounds of this group, independent of the place of their articulation ($p > 0.7$).

In contrast to the aspirational phase, duration of the explosive phase does not depend on the duration of intersegmental pause in the sound sequence ($p > 0.3$). The presence/absence of pause, however, leads to significant differences in burst energy characteristics ($p < 0.001$), namely in maximum intensity of spectral slices, made at points of maximum amplitude deviation of the burst. Thus, with medium pauses maximum intensity amounted to 83.2 dB, while in absence of pause inside the sound sequence – to 74.3 dB. The position of plosives in stressed syllables is not significant for the explosive phase ($p > 0.3$).

On the segment level we observed certain interrelations between the intensity of closure and explosive phase ($r = 0.58$; $p < 0.001$) – the greater the intensity of closure, hence the less tense the closing gesture, the greater the intensity of burst. This correlation may explain the mechanism of lenisation of fortis plosives.

Lack of pause in sound sequences not only leads to reduction of the energy level of sounds, but also to various kinds of sound modifications, one of which is *assimilation as to manner of formation*. In determining the reasons of such sound alterations, we should establish, whether the consonant sequences at word boundaries are homorganic or non-homorganic.

In homorganic sequences plosive+plosive [dt], plosive+fricative [tz], plosive+sonorant [tn] in lack of intersegmental pause the explosive phase of the first sound can merge with the implosive phase of the second sound. In the first case, a homorganic sound sequence transforms into one consonant with maximum duration of the closure phase (up to 100 ms). In the second case, a smooth transition to fissure is formed, thus altering this sound sequence into an affricate. Homorganic sequences of plosives and sonorants are characterized by lack of aspirational and implosive phases, this leading to realization of only the closure and its extension to 70-100 ms. In this case the prejuncture consonant becomes implosive. At the same time, in this type of juncture the oral burst is replaced by a nasal one, which is formed in transition to sonorant due to lowering of the soft palate [10].

In sequences of non-homorganic plosives, while preserving the burst of the prejuncture consonant, the closure of the post-juncture plosive can occur prior to the burst of the prejuncture sound. In this case the order of consonants becomes significant. If the preceding sound in a sequence is a plosive with a deeper place of articulation (e.g., velar [k] before alveolar [t/d] or bilabial [p/b]), its burst is reduced to mere separation of the active organ from the passive one, which is not followed by escape of air. Such burst practically does not have acoustic effect and is characterized by minimum duration (3-5 ms) and weak total intensity (55-60 dB). The level of concentration of burst energy shifts to the low frequency area (1000-1600 Hz) and its intensity does not exceed 75-80 dB. On the spectrogram this is expressed by lack of black-dark gray colors.

A significant increase of closure phase duration and alteration of characteristics of the explosive phase of prejuncture velar plosive [k] in this type of juncture can be explained by sharp diminution of the volume of resonating cavity at the moment of rapid transition of articulating organs from a velar closure to an alveolar or labial one. A high level of anticipating coarticulation of non-homorganic sequences characterizes this type of juncture. Articulatory gestures of labial or alveolar closures are dominant and determine the specific features of velar closure [2; 4]. The effect of anticipation is increased by the degree of polarity (distance) between components of the sequence. Thus, e.g., in sequences of velar and bilabial plosives [k + p/b] the level of anticipation is higher than in sequences of velar and alveolar plosives [k + t/d].

If, however, in a non-homorganic sequence of plosives the prejuncture sound is a plosive with a more front articulation (e.g., [p + k/g], [t + k/g]), it can preserve a short aspirative phase (15-25 ms) only if there is an intersegmental pause of not less than 60 ms. Reduction of duration of the pause inside the sound sequence [t + k/g] leads to the realization of non-aspirated prejuncture plosive [t] with burst intensity level of 85-90 dB and concentration of energy on high frequencies (3000-5000 Hz). This is characteristic of alveolar plosives.

In the realization of plosives an important issue is the strength of articulation, i.e., the level of muscular tension, which can be described in terms of *tense/lax*.

Acoustically, the opposition tense/lax can be characterized by higher general sprectral energy and longer duration, as compared with lower general energy and shorter duration. Such an approach is more useful in the study of vowels, which possess a more clearly expressed formant structure on stationary segments. In contrast, characteristics of consonants consist of both relatively constant and variable values, and this does not enable to determine the formant model of a consonant, which could be a physical correlate of a definite configuration of the speech tract [8].

In the formation of consonants tension develops in one specific organ, i.e. the place, where there is an obstacle, which serves as the noise source of the sound. Such approach is the base for traditional division of plosives into fortes (p, t, k) and lenes (b, d, g), where the tense/lax opposition of consonants takes account only of the strength of the obstacle and air pressure, neglecting the role of vocal cords in pronouncing voiced plosives. In contrast, if we proceed from the total articulation of sounds and take account of the work of vocal cords, then voiced plosives will appear more tense than voiceless ones, e.g. in Slavic languages [9].

The degree of airiness of aspirated plosives is in no way connected with tension of articulatory organs, as friction of air against vocal cords, which occurs at the moment of narrowing of the glottis and accompanies the burst, is formed without tension.

In the present research tension of plosives is primarily connected with tension of the *closing movement*, which is the main articulatory element of all plosives and performs the function of obstacle on the way of the air flux. Therefore, tension of closure can be determined by the share of voice in its formation, which is expressed in terms of increase or decrease of intensity level – the lower the intensity of the closure, the higher its tension. Accordingly, increase of the level of intensity will indicate weakness of the closing movement.

In this paper we also aimed at measuring tension of plosives in terms of acoustic power (squared sound pressure), as it can reflect the degree of voice participation in the articulation of sounds and serve as criterion on their voiceness/voicelessness. Thus, e.g. vowels have maximum values of power (10000-30000 μPa^2); for sonorants this value is somewhat lower (3000-20000 μPa^2); for voiceless sounds theses values vary from 70 to 2000 μPa^2. Increase of power in voiceless consonants can indicate their lenization, and hence, laxation of tension. Therefore, we can conclude that tension of sounds is inversely proportional to their power.

As a result of investigating the relationship between tension of voiceless plosives and their power, we established the following:

1. The duration of aspiration phase in prejuncture plosives has little effect on their power (r = 0.16), as the degree of airiness is not determinative for characterizing sounds as tense.
2. We observed an inverse functional relationship between intensity of the closure phase and its duration (r = -0.71): the longer the duration of the closure phase, the more tense the articulatory closing gesture. Hence, power of plosives, which characterizes the degree of their tension, can be determined by two factors: duration of closure and its intensity. Correlation analysis between duration of the closure phase, its intensity and power of the whole sound showed lack of gradual dependency between these parameters. A certain relationship can be traced between intensity of the closure and power of the sound (r = 0.53), whereas the relationship between power and duration is weaker (r = 0.3). Therefore, the decisive factor in determining the degree of tension of a sound is not the duration its closure, but its intensity.
3. Among the factors effecting laxation of tension of voiceless plosives we should mention interaction of final plosives with preceding sounds. The type of juncture, as well as prosodic factors have less influence on tension of plosives. In cases, when a plosive is preceded by a homorganic sonorant, the closure phase can be significantly reduced or completely absent, as the functions of closure are in this case performed by the sonorant itself (komme**nd** in, Fu**nk** falls).
4. Intensity of the closure phase and, hence, of the whole sound increases if the plosive is preceded by a fricative (Augu**st** von, Geri**cht** Berlin, Wirtscha**ft** wächst). In such combinations the closure phase accounts for 60-70 % of the total sound duration, while its power drops to 100-200 μPa^2.

CONCLUSION

The conducted experimental-phonetic research revealed that final voiceless plosives in consonant sequences at word boundaries are characterized by partial or complete loss of aspiration, reduction of the explosive phase, laxation of tension, this being expressed in partial voicing of voiceless plosives.

The determinative factor in examining interrelations of sounds is the pause pattern inside the sequence. Reduction of pause duration increases mutual influence of the sounds, this leading to elimination of their main phonological characteristics and emergence of various allophonic variants. Sound sequences at word boundaries preserve their quantitative and qualitative characteristics only in presence of maximum intersegmental pauses.

The results of the present research are phonetically interesting and can prove useful in speech culture of mass media, as well as in training accentless pronunciation, logopaedic and phoniatric practice.

REFERENCES

[1] Bukharov, V.M. (1995), Variants of Pronunciation Norm of Modern Literary Language, Nizhny Novgorod. (In Russian)

[2] Fowler, C. & Saltzmann, E. (1993), Coordination and Coarticulation in Speech Production. In: Language and Speech 36, pp.171-195.

[3] Kohler, K. (1977), Einführung in die Phonetik des Deutschen, Berlin.

[4] Kühnert, B. (1997), Die alveolar-velare Assimilation bei Sprechern des Deutschen und Englischen: Kinematische und perzeptive Grundlagen. In: Forschungsberichte, 35, Institut für Phonetik und Sprachliche Kommunikation der Universität München, S.175-379.

[5] Lindner, G. (1975), Der Sprechbewegungsablauf. Eine phonetische Studie des Deutschen, Berlin.

[6] Lotzmann, G. (1975), Zur Aspiration der Explosivae im Deutschen. In: Wiss. Zs. Univ. Halle, Ges. U. Sprachw., Reihe.

[7] Potapova, R.K. & Gordeyeva, T.A. (1998), Towards the Issue of Boundary Signals in Modern German. In: Problems of Linguistics, No.2, pp.118-128. (In Russian)

[8] Potapova, R.K. & Lindner, G. (1991), Characteristics of German Pronunciation, Moscow. (In Russian)

[9] Taranec, V.G. (1997), Energietheorie des Sprechens, Regensburg.

[10] Tillmann, H.G. & Schiel F. (1998), Akustische Phonetik, München.

[11] Zlatoustova, L.V., Potapova, R.K. and Trunin-Donskoy V.N. (1986), General and Applied Phonetics, Moscow. (In Russian)