Phonetic information relevant for spoken word recognition

Yulia Nigmatulina
Department of Phonetics, Faculty of Philology, St.Petersburg State University, Russia
julia.nigmat@yandex.ru

Reduction is an inherent and undeniable feature of spontaneous speech. It can occur inside wordforms as well as across word boundaries. However, lack of phonetic information usually does not prevent successful communication between speakers. Multiple sources of information cooperate at the same moment to carry out the goal of speech recognition. The role of phonetic and phonological information in this process is in the focus of the present research. Special attention is paid to word-external reductions that can modify two words simultaneously.

All two words combinations (W1 and W2) with no pause between the words taken from the Speech Subcorpus of the Corpus of Standard Literary Russian (www.narusco.ru; 65 minutes, or 8786 items overall) were analyzed in the research. The following 6 types of quantitative external reduction were found out: 1) final sound deletion of W1; 2) initial sound deletion of W2; 3) final part (more than one sound) deletion of W1; 4) final sound deletion of W1 and initial sound deletion of W2; 5) sound contraction at word boundary: должны_опираться [doʒne+p'tpʌra+fs] (‘must’-PL_‘base’-INF,MED,); 6) word contraction: страшное_действительно [stra+n sɬi+tma] (‘ugly’-Sg,N_‘actually’). Last two types implicate overlapping of two words. The term ‘sound contraction’ is used for an interaction of only two external sounds when one new sound is pronounced instead of two: while under ‘word contraction’ we mean an interpenetration of two words including more than two sounds.

Two experiments were held to verify how listeners could recognize wordforms in two different cases: 1) if vowel contraction occurs at word boundary; 2) if high external reduction takes place (‘word contraction’). The subjects were asked to write down what they heard using the letters of the Russian alphabet. Word forms with contractions were presented in two conditions: in the phrase context and without it.

The results showed that in the phrase context strongly reduced two-words combinations are recognized worse than the stimuli with only vowel contraction (73% against 98% respectively). Meanwhile, without the context the subjects still reported the existence of two words in the majority of the stimuli, even if they were not able to recognize (identify) the words (the presence of two words was detected if in an answer there was a gap in the letter-string): for example, ‘пиар автмато’ (‘piar avtomaty’* – meaningless letter-sequence in Russian) ‘for the stimulus ’реальная_альтернатива’ [ræ+ɪnatərɪvə] (‘real alternative’-F.SG).

Therefore, the results reveal two questions: 1) what phonetic/phonological information indicate listeners the presence of more than one word, especially when they fail to identify the words; 2) what acoustic information stays unreduced and how listeners could use it to reconstruct proper lexical item.

In order to find relevant features of speech signal that could be crucial for its recognition all answers of the experiment with contracted words as well as the phonetic realizations of all stimuli used were analyzed and compared. Some general principles of information processing were tested based on the sequences of elements reported by all subjects in the answers for each stimulus. Prosodic information and segment characteristics are taken into account. The results of the analysis will be presented and discussed.

1 The work is supported by research grant number №11-1778.2014.6 from the President of the Russian Federation.