Reductions, their perception, and their non-perception

Frank Zimmerer, Goethe University, Frankfurt
zimmerer@em.uni-frankfurt.de

One of the most important features of conversational speech is its variability. The same words are very differently if uttered by two different speakers. Even the same word uttered by the same speaker twice, will not be the same. Human perception seems unimpressed by this variation, however.

On the other hand, we know that even small changes in the speech signal can hamper speech perception processes drastically. A simple change of segments can lead to problems for listeners identifying the correct lexical entries. On a theoretical level, variation, especially due to reduction processes arguably should pose obstacles for speech perception. For instance, the number of segments or syllables can different from the one (canonical) representation due to reduction in conversational speech. Thus, speech perception is not trivial, but works quite well in everyday situations, even when listeners and speakers interact in very noisy surroundings.

Several different approaches have aimed at explaining the data from conversational speech, both with respect to the amount of variation as well as to the ease of listeners in dealing with this variation in speech perception. In this talk I will give an overview over prior research to demonstrate the context in which variation, especially reductions and deletions, is not problematic for perception (and where it is) and, in how far linguistic models have been able to deal with this data.

Broadly speaking, there are two groups of models. They handle variation with two opposing basic assumptions, regarding the number of entries for each word: the first group of models assumes single, abstractionist representations, where variation is not stored, and abstraction processes have to explain how variable input is matched to the entries in the mental lexicon. The second group of models assumes that variation is directly stored in the mental representation via episodic memory entries. In its simplest version, every time a variant of a word is encountered, an extra representation is stored. For the first group of models, the biggest challenge lies in explaining how reductions with drastic and non-rule based deviations from canonical productions can still be perceived correctly. Since models of the second group have an episodic architecture, they assume very detailed and numerous storage of episodes, they do not encounter the obstacle of abstracting from variation. Variation is represented directly in the lexicon. Their main challenge (or the challenge for listeners) can be seen as an activation or search problem. They have to explain, how the correct entry gets activated, and why there is no activation for many other entries that can be rather similar.

In this talk I will present results of several studies where the grade of deviation from the canonical production is varied. Results indicate that contextual information is crucial and that listeners make use of whatever information they can grasp from the speech signal. These results will be evaluated with respect to the assumptions of the different models.