Speech rate and the perceptual restoration of deleted vowels

Background: Spoken language is characterized by reductions, as in (1), so that not all segments of a morpheme are always pronounced (Hooper 1976, Patterson et al. 2003). This presents listeners with a problem. For successful lexical access, listeners need to perceptually restore such unrealized segments. In this paper, we explore some of the factors on which listeners rely when performing perceptual restoration.

(1) support [s_pət] potato [p_təɾəʊt] memory [mɛm_ɹɪ]

When presented with an auditory stimulus like [spət], listeners need to consider at least two perceptual hypotheses, /səpət/ and /spət/. In most instances, the communicative context will supply the information necessary to choose between these hypotheses. However, we show that other sources of information can also bias listeners’ selection towards one of the percepts.

Question: Reductive phenomena, such as the schwa deletion in (1), are more likely at faster than slower speech rates. When presented with a stimulus like [spət], a listener would therefore do better to assume that this stimulus corresponds to /səpət/ if the stimulus is presented at a faster than a slower speech rate. We conducted a perception experiment to investigate whether listeners do in fact use speech rate when performing perceptual restoration.

Design: We recorded 36 non-words of the form [C₁C₂vC₃V] ([pətáki], [kətámi], etc.), and spliced the schwa out of these to create tokens of the form [C₁C₂vC₃V]. Using Praat, we created copies of each token (with and without schwa) at double the rate of the original recording. The full token set (with and without schwa, at original rate and double rate) was then presented to listeners in a syllable-count task.

Hypotheses: (i) Listeners will sometimes perform perceptual restoration in [C₁C₂vC₃V] forms by perceiving a schwa between C₁ and C₂. They will then identify these forms as consisting of three syllables. (ii) If listeners rely on speech rate to determine whether to perform perceptual restoration, they should more often perform restoration at higher speech rates.

Results: Both hypotheses were confirmed – listeners often responded that [C₁C₂vC₃V] tokens consisted of three syllables (evidence for perceptual restoration), and they did so significantly more at the faster (73%) than the slower speech rate (57%) (evidence that they rely on speech rate when performing perceptual restoration, paired sample t(35) = 5.13, p < .001).

Discussion: Listeners make use of all information at their disposal during speech perception. Any systematic pattern in the speech signal can be exploited. This study shows that listeners have knowledge of the systematicity with regard to schwa deletion at different speech rates, and that they use this knowledge during perception. More frequent words are also subject to more frequent reduction (Hooper 1976; Patterson et al. 2003). Similarly, more frequent words are produced faster (Gahl 2008). We therefore hypothesize (and are exploring in ongoing research) that listeners will be more likely to perform schwa restoration in frequent than infrequent words, and that this might be at least partially explained by the result found here, namely that schwa restoration is more likely at faster than slower speech rates.
References

