Beginning to understand how intensity influences speech perception

John Kingston[1]

[1]*University of Massachusetts, Amherst (USA)*
jkingstn@umass.edu

We might expect a more intense sound would simply provide more energy, and thereby convey greater prominence, and possibly that the sound is or belongs to a stronger prosodic constituent. Or greater intensity might trade off perceptually with a longer duration, as Repp (1979) showed for aspiration; for a fuller account of this trade off among the correlates of [voice] judgments, come to our ICPhS poster in Prague. The results to be reported there, like those that are the focus of this talk, show how the relative intensity of one acoustic interval does and does *not* influence the perception of a neighboring interval. While the specific purpose of the experiments reported here is to provide evidence relevant to the debate about whether the objects of speech perception are articulatory gestures or auditory qualities, their more general purpose is to remedy the neglect of a ubiquitous acoustic property.

All the experiments examine how the intensity of a preceding /al/ or /ar/ context influences the categorization of a following /da–ga/ target continuum. Since Mann (1980) first reported that listeners respond "ga" more often after /al/ than /ar/, manipulations of such stimuli have been used to argue that the objects of speech perception are the articulatory gestures that produce the speech signal's acoustic properties, or alternatively, that they are the auditory qualities evoked by those acoustic properties (for reviews, see Fowler, 2006; Lotto & Holt, 2006). Listeners also respond "ga" more often after a non-speech analogue of /al/, which suggests that they perceive the target as contrasting spectrally with that spectrally high context, rather than compensating for coarticulation with that more anterior context. Three experiments test and reject the alternative mechanism, informational masking, which Viswanathan, Fowler, & Magnuson (2009); Viswanathan, Magnuson, & Fowler (2013) have proposed as responsible for the shift in categorization produced by the non-speech contexts. The first does so by showing that more intense speech contexts do not increase the size of the shift, the second by showing that more intense non-speech contexts don't do so either, and that the shift gets smaller as the spectral distance between the contexts and targets increases, and third by showing that non-speech contexts with energy in the same auditory bands as the target shift categorization less than those with energy in complementary auditory bands. If they are available in time, the results will be discussed of a fourth experiment that tests an alternative account of Viswanathan et al.'s (2009) finding that band-passed speech contexts which preserved the spectral difference between /al/ and /ar/ but which were thereby rendered non-speech failed to shift categorization.

Taken together, these results show that the target's perception is not simply determined by how intense its context is, but by how energy is distributed across the target's and context's spectra; that is, by how intensity is greater at some frequencies than others, and how differences between target and context in which of their frequencies are more intense interact perceptually within and between target and context. Just as we have learned that some milliseconds influence perception more than others, so, too, do some deciBels. (These studies have all been carried out in collaboration with Amanda Rysling, of the University of California, Santa Cruz.)

References

Fowler, C. A. (2006). Compensation for coarticulation reflects gesture perception, not spectral contrast. Perception and Psychophysics, 68 , 161-177.

Lotto, A. J., & Holt, L. L. (2006). Putting phonetic context effects into context: A commentary on Fowler (2006). Perception and Psychophysics, 68 , 178-183.

Mann, V. A. (1980). Influence of preceding liquid on stop-consonant perception. Perception & Psychophysics, 28 , 407-412.

Repp, B. H. (1979). Relative amplitude of aspiration noise as a voicing cue for syllable-initial stop consonants. Language and Speech, 22 , 173-189.

Viswanathan, N., Fowler, C. A., & Magnuson, J. S. (2009). A critical examination of the spectral contrast account of compensation for coarticulation. Psychonomic Bulletin and Review , 16 , 74-79.

Viswanathan, N., Magnuson, J. S., & Fowler, C. A. (2013). Similar response patterns do not imply identical origins: An energetic masking account of nonspeech effects in compensation for coarticulation. Journal of Experimental Psychology: Human Perception and Performance, 39 (4), 1181-1192.