Perception of voicing is graded all the way down: Evidence from electrophysiology

Joseph C. Toscano | Bob McMurray
University of Iowa, Dept. of Psychology
joseph-toscano@uiowa.edu | bob-mcmurray@uiowa.edu

Phonological distinctions, such as voicing, are manifested as variations in continuous acoustic cues in the speech signal. These cues are ultimately categorized into a small number of distinct phonological categories. At what levels of processing are graded, continuous representations available? Recent experimental evidence has shown that listeners are sensitive not only to changes in phonemic category, but also to within-category differences, as evidenced by eye-tracking (McMurray, Tanenhaus, & Aslin, 2002) and category goodness ratings (Miller & Volaitis, 1989). However, these measures tap relatively late stages of processing that correspond to category levels of representation (i.e. fixation probabilities reflect lexical access and goodness ratings reflect category judgments). Until recently, there has been no way to measure listeners' responses to continuous cue-values at stages that precede categorization. The present study asks whether we can simultaneously observe within-category sensitivity early in speech perception at the level of both cue and category. Does sensitivity at late stages arise from similar sensitivity at the level of continuous acoustic dimensions?

We conducted an event-related potential (ERP) experiment designed to elicit responses to both early encoding of auditory information and phonemic categorization. Listeners categorized spoken words (*beach*, *peach*, *dart*, and *tart*) that varied in voice onset time (VOT) in an auditory oddball task. We found two ERP components that reflected variation in VOT. The P3 component (a measure of categorization) varied in amplitude as a function of distance from the target phonetic category (Figures 1a and 2a). Thus, this component likely reflects the gradient activation of words or categories. A second ERP component, the auditory N1, varied in amplitude and timing as a function of VOT, but not as a function of the target category (Figures 1b and 2b). In addition, we also found a small modulation in the N1 as a function of place of articulation. Thus, this component reflects processing at the level of continuous acoustic cues, not categories. These results suggest that the speech system is sensitive to within-category differences in VOT along both acoustic feature and phonological category dimensions.

These results provide further evidence that listeners are extremely sensitive to fine-grained acoustic detail, and they suggest that this detail is maintained from the earliest stages of perceptual processing through categorization. Furthermore, these results do not indicate a point at which discrete representations of phonemes are observed in speech processing, which may have implications for phonological theories that posit such discrete representations.

References

McMurray, B. Tanenhaus, M. K., & Aslin, R. N. (2002). Gradient effects of within-category phonetic variation on lexical access, *Cognition, 86*, B33-B42.

Miller, J. L. & Volaitis, L. E. (1989). Effect of speaking rate on the perceptual structure of a phonetic category. *Perception & Psychophysics, 46*, 505-512.

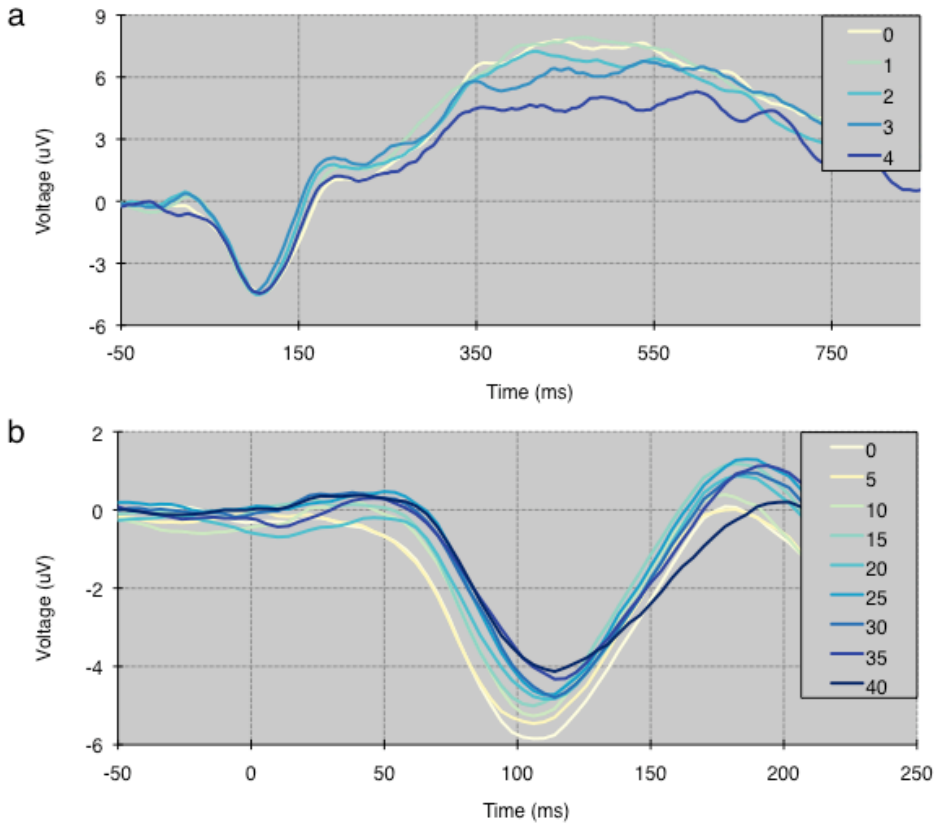

Figure 1. (a) P3 recorded from parietal channels (P3, Pz, P4) for conditions as a function of distance (in 5 ms VOT steps) from the target VOT (0 = target item; target responses only). (b) Auditory N1 recorded from frontal channels (F3, Fz, F4) for each VOT.

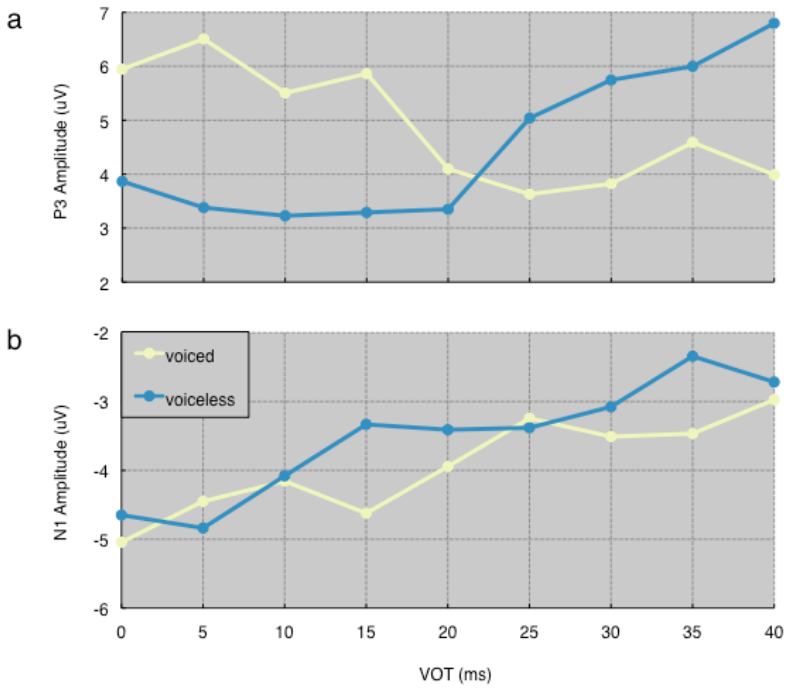

Figure 2. (a) Mean amplitude of P3 component for each VOT. Amplitude increases as VOT approaches the endpoint of the target category. (b) Mean amplitude of N1 component for each VOT. Amplitude decreases as VOT increases for both voiced and voiceless targets.