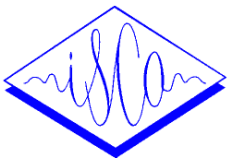

ITRW on Temporal Integration in
the Perception of Speech
Aix-en-Provence, France
April 8-10, 2002

P3-12

Gradient effects of within-category VOT differences on lexical activation as measured by eye fixations

Bob McMurray[1], Michael K. Tanenhaus[1], Richard N. Aslin[1], and Michael J. Spivey[2]
1: Department of Brain and Cognitive Sciences, University of Rochester, NY, USA
2: Department of Psychology, Cornell University, NY, USA
mcmurray@bcs.rochester.edu, mtan@bcs.rochester.edu, aslin@cvs.rochester.edu, spivey@cornell.edu

Although small differences in within-category acoustic information have limited effects on explicit phonemic decisions, we present evidence that they have strong effects on the magnitude and time course of activation for lexical alternatives. On each trial subjects heard either a token from a 9-step synthetic /b/-/p/ VOT continuum embedded in a word, or an /l/- or /ʃ/-initial filler word. They responded by clicking with a mouse on one of four simultaneously presented pictures on each trial: a picture for both the /b/ and /p/ response, an /l/- filler item and an /ʃ/- one (e.g. *peach*, *beach*, *leaf* and *ship*). We used eye movements as the dependent measure, building on recent studies that have established a clear link between activation levels for lexical candidates in models of word recognition and proportion of fixations.

We analyzed the proportion of fixations to the 'target' (e.g., *beach* when a 0 ms VOT token was presented) and the proportion of fixations to that token's most similar alternative or 'competitor' (e.g., *peach*). We examined prototypical VOTs (those that had 95% or more responses to one category): 0, 5 and 10 ms for /b/; 25, 30, 35 and 40 ms for /p/. We excluded any trial in which the 'incorrect' word for that category was selected (by the mouse click). As VOT neared the category boundary, the proportion (and duration) of fixations to the competitor picture increased, resulting in a significant effect of within-category VOT ($F_{ba}(2,32) = 6.92$, $p = 0.003$; $F_{pa}(3,48) = 7.23$, $p < 0.001$). Significant linear trends were found even when we excluded the VOTs nearest to the category boundary and we separated the trials into early and late fixations. Thus fine-grained within-category differences in VOT are preserved in patterns of lexical activation. We argue that this graded activation could be used probabilistically to help listeners resolve local temporal ambiguity and integrate acoustic/phonetic features across time.

We also used a similarly constructed VOT continuum in a non-word CV context. In separate experiments, subjects made either a 2AFC or 4AFC phoneme decision by clicking on a 'P' or 'B' button or one of four buttons ('P', 'B', 'L' and 'Sh'). In contrast to the word identification task, the 2AFC and 4AFC tasks showed no effect of VOT within prototypical ranges of the continuum on looks to the competitor button (2AFC: $p_{ba} > 0.1$, $p_{pa} > 0.1$; 4AFC: $p_{ba} > 0.1$, $p_{pa} > 0.1$). However, the proportion and duration of looks to the competitor did increase at the VOTs immediately adjacent to the category boundary ($t(31) = 2.854$, $p = 0.008$).

Analyses of the identification data from the mouse clicks showed a large effect of task on the slopes of the identification functions ($F(2,47) = 31.41$, $p < 0.0001$). The 2AFC data showed the steepest slope, followed by the 4AFC phoneme decision data, and then the 4AFC word data. This suggests that a simple shift in the task from 2AFC to 4AFC increased reliance on detailed acoustic/phonetic information, which was further increased when the information was embedded in words. Thus demonstrations of poor discrimination of within category differences in 2AFC phoneme categorization and discrimination tasks underestimate the degree to which fine-grained within category differences might affect real-time speech perception. [Supported in part by NIH Grants NIDCD DC-005071 and NIDCD T30 DC00035.]