The Use of Metrical Prosody to Segment Continuous Speech

Kathleen Eberhard (KATHLEEN.M.EBERHARD.1@ND.EDU)
Carol Madden (CAROL.J.MADDEN.20@ND.EDU) **Kerry Zahn** (KERRY.A.ZAHN.2@ND.EDU)
Department of Psychology, University of Notre Dame
Notre Dame, IN 46556 USA

Introduction

This study investigated the processes involved in mapping a spoken signal onto lexical-form representations stored in the mental lexicon. Current models of word recognition assume that the process begins when the initial portion of a word's acoustic-phonetic input makes contact with, or activates all corresponding items in the lexicon. Items drop out of the activated set when they are no longer consistent, and recognition occurs when a unique item remains. Prior research shows that recognition time is influenced by factors such as the number of competing items, their frequency of occurrence, and whether contextual information is available to help eliminate items.

Recently, research has begun to focus on questions about the nature of the contact representation. Following work in current phonological theory, Lahiri and Marslen-Wilson (1991; also Marlsen-Wilson & Warren, 1994) have proposed that the lexical representations consist of a featural array containing only marked features. Support for this proposal comes from studies employing two of the few techniques available: gating and auditory lexical decision. The goal of the work here was to use a new technique to examine the competition among activated items that differ in the markedness of their feature specifications.

Methodology

The technique was recently developed by Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy (1995), and involves recording listeners' eye movements as they follow instructions to move objects in a display (e.g., "Pick up the candy"). Eye movements are tracked using (ASL, Bedford MA) an eye camera attached to a helmet. The helmet also contains a video camera that displays the listeners' field of view on a TV along with his/her eye-position record. A VCR synchronously tapes the video record and the spoken instructions. Data are analyzed using (33 msec) incremental playback of the tapes to locate word onsets and measure eye movement latencies occurring with respect to them.

Tanenhaus et al (1995) found that listeners' eye movements are closely time-locked to the acoustic input, making them a sensitive measure of the word recognition process For instance, when told to "pick up the candle", listeners looked at the candle shortly after the end of the word. When the instruction was given in a context containing both a candle and candy, the looks were delayed by 30 msec.

The current study used a similar procedure to examine the competition between words that contrast in whether their initial syllable has a full (marked) or reduced (unmarked) vowel. Recorded instructions such as, "Pick up the bullet" (marked) or "Pick up the balloon" (unmarked) were presented in competitor-present contexts (bullet and balloon present), as well as in competitor-absent contexts, (no related object names). Eye-movement times were compared to examine whether "unmarked" competitors delay recognition of "marked" targets and vice versa.

Results

Fig. 1 shows data from 4 subjects. Analyses showed that unmarked competitors reliably delayed recognition for marked targets ($t(3)=3.75$, $p<.02$), but marked competitors did not delay recognition of unmarked targets ($t(3)=1.18$). These results suggest that both words with full and reduced forms of the same vowel are activated by the full vowel's marked specifications, resulting in competition. However, when no marked specifications occur in the signal, only words with the reduced vowels are activated.

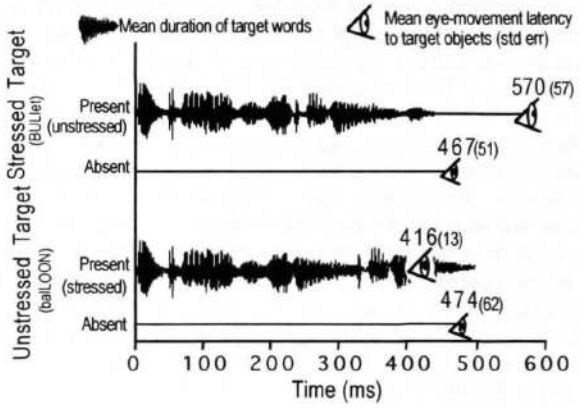

Figure 1: Results of Experiment 1.

References

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632-1634.

Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition, 38*, 245-294.

Marslen-Wilson, W., & Warren, P. (1994). Levels of perceptual representation and proces in lexical access: Words, phonemes, and features. *Psychological Review, 101*, 653-675.