Beyond Monosyllables: Word Length and Spoken Word Recognition

Ted Strauss (ted.strauss@gmail.com)
Department of Psychology, New School for Social Research and University of Connecticut
James S. Magnuson (james.magnuson@uconn.edu)
Department of Psychology, University of Connecticut and Haskins Laboratories

Abstract

Most empirical work in spoken word recognition has focused on monosyllabic words, and most theories and models are limited to monosyllables either in scope or in practice (in that the theories' predictions for polysyllabic words have largely been unexplored). We add to the small number of studies to date explicitly comparing recognition of mono- and polysyllabic words with a study examining time course predictions that follow logically or by simulation from three theoretical perspectives. We tracked subjects' eyes as they followed spoken instructions to click on pictures in a display (i.e., we used the visual world paradigm). Consistent with a computational model, we found a late-emerging advantage for long words over short words matched on several lexical dimensions, though the effect was weaker than expected. The results suggest new constraints on theories of spoken word recognition.

Keywords: spoken word recognition; lexical activation; word length; eye tracking.

Introduction

Most of what we know about how spoken words are recognized is based on studies using monosyllabic words. And most theories and models of spoken word recognition (SWR) are limited (in scope or in practice, in terms of what has been explored) to monosyllabic words. However, there is growing evidence that short and long words may be processed differently, or at least that the constraints that determine processing facility depend on word length. We will briefly review how two major theories apply primarily to monosyllables, and predictions that would follow if they were extended to polysyllabic words. We then compare those predictions, as well as predictions from simulations of the TRACE model (McClelland & Elman, 1986), with time course data from a study comparing recognition of long and short words.

Long words and theories of SWR

Current theories agree on some basic facts regarding SWR: as a spoken word is heard, multiple lexical items are activated as a function of similarity to the input and prior probability (word frequency), and activated items compete for recognition. A primary point of difference between theories is the similarity metric each assumes, which determine the predicted competitor set. Similarity metrics have primarily been based on empirical work with monosyllabic words. In this section, we will review two theories with explicit similarity metrics (the Neighborhood Activation Model [Luce & Pisoni, 1998] and the Cohort model [Marslen-Wilson & Welsh, 1978]) and an implemented model (TRACE) with an "emergent" similarity metric, and predictions that follow from each view for long vs. short words.

The **Neighborhood Activation Model** (Luce & Pisoni, 1998) uses a neighbor-based similarity metric, and predicts that how easily a word is recognized depends on its *neighborhood density* (Luce & Pisoni, 1998): the frequency-weighted count of its *neighbors,* most commonly defined as any word differing from a target word by a single phonemic deletion, addition, or substitution. Note that this is a global similarity metric; what matters is how much overlap there is between two words, not "when" the overlap occurs. When a target word is heard, its neighbors are assumed to be activated and to compete for recognition. Thus, the more neighbors a word has, the slower its recognition is predicted to be.

This metric is essentially undefined for polysyllabic words; while Luce and colleagues have reported exploratory investigations of how the neighborhood metric can be extended to longer words (e.g., Cluff & Luce, 1990; Vitevitch & Luce, 1999), the model is thus far limited to monosyllabic and (to a limited degree) bisyllabic words. However, we can examine what predictions follow if we apply the familiar 1-phoneme neighbor threshold to polysyllabic words. If we do this, an interesting prediction for long words can be inferred. As word length increases, the likelihood of single-phoneme neighbors decreases (e.g., *cat* has many neighbors, such as *cab, at, cot,* and *bat*, while *catapult* has none). This leads to an overall negative correlation between word length and neighborhood density in the lexicon, $r = -0.44$ (all lexical correlations we report are based on analyses of a 27,000 item lexicon compiled from the American National Corpus, Ide & Suderman, 2003). Thus, (ignoring durational differences for now) the 1-phoneme neighbor metric predicts that long words should be recognized more quickly than short words, since they will typically have many fewer neighbors to compete with (though as we discuss below, the time course of lexical activation and competition depends on the temporal distribution of similarity over words; Magnuson, Dixon, Tanenhaus, & Aslin, 2007).

In contrast to the global neighborhood metric, the **Cohort model** (Marslen-Wilson & Welsh, 1978) assumes that the effect of overlap depends on the temporal distribution of similarity. Specifically, it assumes that words are deactivated when they mismatch the bottom-up input. So as you hear the word *cat*, initially all words beginning with /k/ are activated, forming an activation *cohort*. When you hear /æ/, all words that mismatch drop out. This means the active competitor set changes over time. Thus, as with the Neighborhood Activation Model, we can predict that the more cohorts there are (with some minimal amount of overlap, such as the first two phonemes), the slower recognition will be. But the theory also specifies how the cohort is winnowed down (reducing potential competition) as more of the word is heard, until the *uniqueness point* (Grosjean, 1980) is reached, where only a single match is possible (e.g., *p* in catapult). Thus, how quickly a word is recognized is predicted to depend not just on how many words overlap with it at onset, but by how large the cohort is, segment-by-segment, over the length of the word.

This leads, potentially, to a prediction of long word disadvantages, as long words tend to have later occurring uniqueness points. For example, *catalog* has both *cat* and *cattle* initially embedded in it, so it should take longer to uniquely identify than *cattle*, which has only *cat* embedded. Indeed, over the entire lexicon, the correlation of word length and uniqueness point in American English is 0.44. We must qualify this by noting that for many short words, the uniqueness point is after word offset. For example, *cat* is not unique at word offset. Distinguishing *cat* from longer words in its cohort depends on hearing silence after *cat* or a word onset (or a longer portion of a word) that is incompatible with any item in the cohort. Thus, the long word disadvantage prediction may only apply to isolated words, or utterance-final words (as we use below).

Unlike the two models discussed so far, **TRACE** (McClelland & Elman, 1986) does not include an explicit similarity metric. TRACE is an implemented connectionist model that maps pseudo-spectral features to phonemes and phonemes to words. There is also feedback from words to phonemes, and lateral inhibition at the phoneme and word levels. Interestingly, the "emergent" similarity metric of TRACE is intermediary between neighborhoods and cohorts. The degree to which words compete with one another depends on the processing dynamics of the implemented model. Words are activated as a function of their bottom-up match to the input. Because the model does not explicitly track word onsets, a word can become activated at any time if it is sufficiently similar to the input. However, activated words inhibit other words. The result is that words like rhymes, which have high global similarity to a target word despite initial mismatch, are predicted to compete (consistent with a neighborhood metric, but not with a cohort metric). However, because activated words inhibit other words, there is an advantage for initial overlap – items activated early on get to inhibit words that initially mismatch the input. For example, Allopenna et al. (1998) reported simulations showing that when a word like *beaker* is presented, *beaker* and its cohorts quickly become highly activated. As the input diverges from the cohorts and becomes similar to a rhyme (*speaker*), the rhyme becomes activated. But it never becomes as active as cohort items because it is inhibited by the target and its onset cohorts.

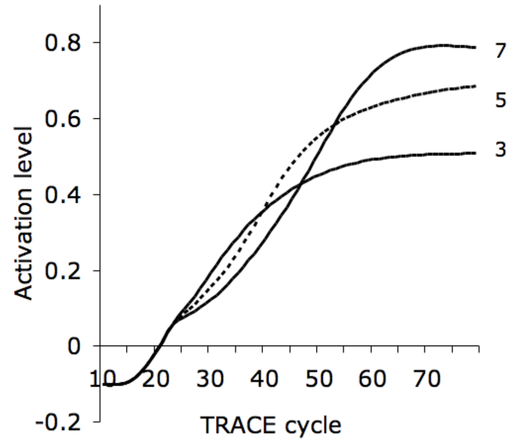

Figure 1. TRACE simulations of 3, 5 and 7 segment words, based on Frauenfelder and Peeters (1990). TRACE predicts an initial long word disadvantage, and later long word advantage.

It is possible to include mono- and polysyllabic words in TRACE's lexicon, so it is easy to compare TRACE predictions for long and short words. In Figure 1, we present results from new simulations conducted using jTRACE (Strauss et al., 2007) based on simulations reported by Frauenfelder and Peeters (1990). TRACE's processing dynamics predict a two-part length effect. Early in processing, long words are more susceptible to the influence of lexical competition than short words (essentially because they receive more inhibition because they overlap in time with more possible words; see the discussion). Later in processing, long words benefit from greater total bottom-up support than short words. Together, these aspects of TRACE predict an initial disadvantage and later advantage for long words. Long words also receive the benefit predicted by the neighborhood model: by the end of a long word, there are few words in the lexicon that are similar with the full-length of the word, leading to less competition at word offset than there would typically be for a short word.

In summary, we have 3 distinct predictions about long words. The Neighborhood Activation Model suggests a long word advantage, Cohort suggests a long word disadvantage, and TRACE predicts an early disadvantage followed by a late advantage. Before testing those predictions, we will review what is already known about differences in processing polysyllabic words.

Processing long words

There has been fairly little empirical work focused on long vs. short words. Grosjean (1980) documented that as word length increases, a larger portion of the word is required to support recognition. This does not imply that long words take longer to recognize simply because they have longer duration, as the proportion of the word required depends explicitly on uniqueness point. Furthermore, phonetic information tends to be somewhat compressed in long words, leading to modest differences in average duration between long and short words. For example, the same syllable (e.g., *cat*) has longer duration when it occurs as a monosyllabic word rather than as part of a polysyllabic word (Lehiste, 1972).

It has only recently been discovered that spoken word recognition is sensitive to such differences: there is stronger priming between cohorts that are similar in length than between a short/long cohort pair (Davis et al., 2002), and there is time course evidence that cohorts matched in length compete more strongly (Salverda et al., 2003). While these findings do not speak directly to the predictions we have just reviewed, they do emphasize the need to learn more about the time course of processing long vs. short words.

A study by Pitt and Samuel (2006) is more pertinent. Employing the Ganong paradigm (Ganong, 1980), they tested predictions that follow from the overall activation advantage predicted for long words in a model like TRACE. They predicted that lexical effects on phoneme decisions should be stronger in long words than in short words because long words would provide greater lexical feedback to constituent phonemes due to their greater overall activation. The results from a series of phoneme categorization experiments confirmed that long words were associated with larger lexical effects, which also took longer to decay than lexical effects with short words. This result is consistent with TRACE's prediction of an eventual long-word advantage, but does not allow a test of the more specific prediction that the advantage emerges late, since they did not employ a time course measure.

The study we present here uses the "visual world paradigm" (Tanenhaus et al., 1995) to examine time course. In this paradigm, we track participants' eye movements as they follow spoken instructions to perform a visually guided task (clicking on objects in a display). The paradigm has been used to examine time course predictions regarding the makeup of the competitor set (Allopenna et al., 1998), the temporal locus of frequency effects (Dahan et al., 2001a), and sensitivity to subcategorical cues (Dahan et al., 2001b). For discussion about how this paradigm indexes lexical activation, and how to link time course estimates from eye tracking to time course predictions from a model like TRACE, see any of these papers.

As a first step in elucidating the time course of lexical activation as a function of word length, we will present results from a simple eye tracking experiment estimating the time course of lexical activation for short and long words.

Experiment

We examined the time course of lexical activation for short and long words simply by measuring how quickly target fixation proportions increased from word onset. We anticipated that differences between long and short word trajectories might be subtle. To increase the likelihood of detecting differences, we ran two groups of subjects. For one, long and short targets were always displayed among unrelated distractors (no-competitor condition). For the other, critical trials included the target, a phonological competitor, and two unrelated distractors (competitor condition). The motivation was that competitor presence might prevent ceiling effects for easily recognized targets.

Method

Participants. 41 students from the University of Connecticut participated for course credit. 21 took part in the non-competitor condition, 20 in competitor condition.

Materials. The 2 versions of the experiment included the same 120 target words and 8 practice trials. There were 36 critical trials (18 monosyllable and 18 trisyllable targets). In the no competitor condition, target pictures were presented with pictures of 3 phonologically and semantically unrelated items. In the competitor condition, the same targets were presented with 1 length-matched onset competitor and 2 distractors. Filler trials included a mix of short and long words. Trials were randomized by subject.

Several characteristics of our short and long words are detailed in Table 1, based on analyses of the American National Corpus (ANC; Ide & Suderman, 2004). Because our focus is on differences in processing due to word length, lexical variables that co-vary with length were not controlled across conditions. Rather, items were selected to be typical of their length. For example, neighborhood density (r=-0.44) necessarily correlates with word length. Neighborhood density is defined as the frequency weighted sum of items that differ from the target by one segment (Luce & Pisoni, 1998). The question of how to reduce the length biases inherent in this metric has been explored in some detail (Storkel, 2004), with no clear solution. Short and long words also differ in duration. Matching on duration would require different speaking rates, so we opted for natural variation in this dimension.

In the competitor condition, onset competitors matched at least the first two phonemes of the target item, e.g., *bomb-box*. Mono- and polysyllabic target-competitor pairs were roughly matched on log frequency per million, imagibility, and neighborhood density (all t-tests N.S.).

Pictures were color photographs, realistic graphics, and some simpler graphic representations, compiled from a variety of sources. Distractor items were matched pseudo-randomly with constraints to keep mean distractor frequency fairly constant and to avoid using phonologically or

Table 1: Stimulus statistics. 18 items per condition. Cohort density was defined as the frequency weighted sum of items that matched the onset of the item's initial consonant-vowel (cf. Magnuson et al., 2007).

Syl	Length	UP	Log freq	Imagibility
1	3.0	2.8	1.21	585
3	7.1	5.7	0.98	583

Syl	Duration (ms)	Cohort density	Neighborhood density
1	465	44.17	27.72
3	602	44.88	12.40

semantically related distractors. Picture positions were randomized by trial. Digital audio recordings were made in a sound attenuated booth by the first author.

Procedure. Each trial began with a fixation cross displayed in the center of the screen. When subjects clicked on the fixation cross with the computer mouse, 4 pictures appeared in the corners of the screen. A 500 ms preview was followed by presentation of the stimulus word over headphones. Subjects were instructed to click on the picture matching the word. After clicking, pictures disappeared and there was a two second pause before the fixation cross reappeared to start the next trial. Gaze position and duration were recorded using an ASL 6000 eye-tracker. Fixation analyses were performed over the raw gaze data using ILAB software (Gitelman, 2002).

Results

Accuracy on the task was at or above 98% for all subjects. Because ANOVAs revealed no reliable main effects or interactions in accuracy or response time, we will focus on the time course data. Problems with the eye-tracker rendered 4 and 5 subjects' data unusable in the competitor and no competitor versions, respectively. Thus, our gaze analysis includes data from 16 subjects per competitor condition.

Fixation proportions over time to targets, competitors and distractors, relative to the onset of the auditory stimuli, are shown in Figures 2 and 3. Fixations were summed across 100 millisecond time bins in figures, and over 10 ms bins for growth-curve analyses. Typical of "visual world" experiments, fixation proportions to target items begin to diverge from chance about 200-300 ms after stimulus onset and rapidly increase towards peak probability of around 0.9. While there is a weak trend towards a trisyllabic advantage in Figure 2, there is a stronger trend in Figure 3. This data presents significant statistical challenges. One common approach is to break the time course into multiple time bins and to include bin as a factor in a repeated measures ANOVA. There are several problem with this approach, including violations of independence assumptions, increased Type I error, and possibilities of arbitrary window sizes leading to spurious results or masking real differences (see Mirman et al., in press). Thus, we opted to use growth curve analyses ("GCA"; Magnuson et al., 2007; Mirman et al., in press; Singer & Willet, 2003). GCA is a multi-level regression approach designed to formally model variations in over-time trajectories. Conceptually, it is analogous to fitting curves (using orthogonal power polynomials; see Mirman et al.) to the time course data for each individual and then analyzing curve parameters (such as intercept and slope) as a function of experimental conditions. This approach avoids the many problems associated with common ANOVA approaches to visual world data.

We were particularly interested in effects on intercept and slope in a model including length and competitor presence factors. In the power polynomial approach, intercept is centered in the curve. Thus, an effect of intercept indicates whether the curve is shifted up as a function of condition, analogous to a main effect of mean fixation proportion in ANOVA. An effect of "slope" (linear change over the entire trajectory) indicates that the trajectories change at different rates, which would be analogous to an interaction of condition and time (because differences in rate of change would lead to differences of varying magnitude at different points in time in a variable like mean fixation proportion). So for example, a significant effect of slope would indicate that the late trend towards a trisyllabic advantage in Figure 3 is significant (i.e., that length interacts with "time" [early vs. late] in ANOVA terms).

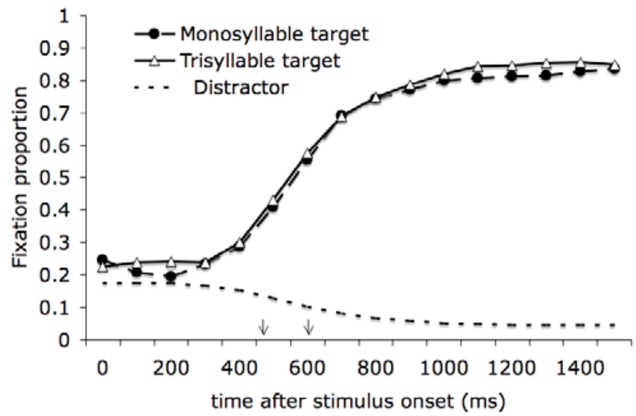

Figure 2. Proportion of fixations to monosyllable pictures and trisyllable pictures in the no competitor condition. The first arrow indicates monosyllable stimulus offset (465ms) and the second arrow indicates trisyllable offset (602ms).

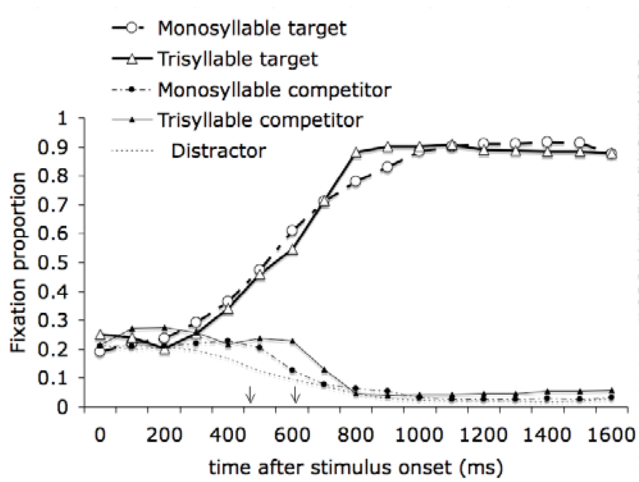

Figure 3. Proportion of fixations to monosyllable and trisyllable target and competitor pictures in the competitor condition.

The growth curve analyses included data from the range of 150 to 1050 ms, covering the theoretically earliest possible moments of signal driven fixations up until peak fixation proportions. In the no-competitor condition, the effect of length was not significant in either the intercept or linear terms. Figure 3, the competitor condition, suggests a cross-over effect in which a slight monosyllable advantage turns into a trisyllable advantage at some point between 600-800 ms. Word length did not have a reliable effect on the intercept, $B = 0.0042$, $SE = .015$, $\Delta D(1) = .027$, *ns* (indicating mean fixation proportion to monosyllabic and trisyllabic targets did not differ statistically). There was a reliable effect on slope, $B = 0.281$, $SE = .16$, $\Delta D(1) = 2.656$, $p = .008$. As mentioned above, this is analogous to an ANOVA interaction between length and time (see Magnuson et al., 2007, for a similar cross-over analysis).

Discussion

This study presents the first time course data explicitly examining differences in the dynamics of the recognition of long and short spoken words. We prefaced the study by examining predictions that follow from three theoretical perspectives on spoken word recognition. While the Neighborhood Activation Model (Luce & Pisoni, 1998) has not been elaborated in detail for polysyllabic words, if the monosyllabic version is applied to polysyllabic words, it predicts a long word advantage because long words will have sparser neighborhoods. We discussed how the Cohort model (Marslen-Wilson & Welsh, 1978) predicts a disadvantage for long words since uniqueness point is correlated with word length. Finally, we presented predictions from TRACE (McClelland & Elman, 1986) in the form of simulated activation time course for words of varying lengths; TRACE predicts an initial early advantage for short words (consistent with Cohort, though based on competition dynamics) followed by a late advantage for long words (consistent with the Neighborhood Activation Model, but based both on greater overall bottom-up input for long words and for sparse competition late in the time course of long words).

Our results were most (though only partially) consistent with the TRACE predictions, but only when a phonological competitor was included in the display. While some competition effects have previously been shown to be enhanced by competitor presence (Dahan et al., 2001b), this suggests that the differences between the monosyllabic and trisyllabic targets used in our study were fairly subtle. Nonetheless, when a competitor was present, we found what conceptually amounts to an interaction with time: an effect on slope in a growth curve analysis comparing long and short targets, indicating a cross-over from a short advantage to a long advantage (though clearly the late long advantage was much greater). Note also that our analyses compare the time course data directly for the two target lengths, without consideration of duration differences for long and short words. Were we, for example, to normalize the time course based on mean msecs per phoneme, the length effect would be strengthened. We have opted against such normalization out of concern for possible distortions it might introduce.

It is worth considering in detail how TRACE arrives at its time course predictions. TRACE's processing dynamics account for these effects via a sequence of processing steps. Initially, weakly activated long words are more susceptible to lexical competition because they overlap temporally with a larger number of words. In TRACE, words receive inhibition from all words with which they overlap. Thus, longer words have more inhibition sites. Next, long words benefit from a stronger bottom-up signal contained in the longer duration of the input, permitting the long word to make rapid gains in activation. Finally, towards the peak of activation, long words benefit from smaller neighborhoods, a competition effect approximating NAM's prediction.

Thus, understanding how and why there may be functional differences in the processing of short and long words requires time course measurements and consideration of the moment-to-moment makeup of the activated competitor set. That is, it is crucial to consider how competition waxes and wanes as a function of temporary bumps in similarity that result in denser *regions* of a word (Vitevitch, 2002; Magnuson et al., 2007), as well as consideration of how contiguity of similarity between input and a lexical representation protects a lexical item from competition. This view would predict that polysyllable word neighborhood densities are concentrated in the first syllable, with very sparse neighborhoods later on. Thus, the majority of lexical competition for long words occurs early in processing, while activation is largely free to accelerate as later portions of long words are heard.

Another interpretation of the initial long word disadvantage focuses on bottom-up processing. During the earliest phase of processing, differences as a function of word length may be due more to information density of the signal rather than the competition dynamics that are the basis for the early differences predicted by TRACE. As we mentioned earlier, segments in long words tend to be compressed relative to segments in short words (Lehiste, 1972), but this fact is not included in typical TRACE simulations (although it could be). Segment compression entails that more acoustic information must be mapped to the lexicon during the same stretch of time for long vs. short words. This raises the possibility of a bottleneck effect early in processing. This interpretation, distinct from the other three models, is being explored in ongoing research.

In conclusion, the current results provide the first time course estimates of differences in the processing of short and long words, and add to demonstrations that the time course of lexical activation and competition cannot be captured by static competitor set definitions. Rather, "competition density" is a factor of multiple constraints. When word length is added to the mix of constraints, an acceleration of lexical activation for long words occurs late in the time course of bottom-up processing.

Acknowledgments

This research was supported by National Institute of Deafness and Other Communication Disorders Grant DC-005765 to James S. Magnuson and National Institute of Child Health and Human Development Grant HD- 01994 to Haskins Laboratories. Ted Strauss was supported by a PGS-

M scholarship from the Natural Sciences and Engineering Resaerch Council of Canada.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*(4), 419-439.

Cluff, M.S. & Luce, P.A. (1990). Similarity neighborhoods of spoken two-syllable words: Retroactive effects on multiple activation. *Journal of Experimental Psychology: Human Perception & Performance, 16*, 551-563.

Dahan, D., Magnuson, J. S., & Tanenhaus,M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology, 42,* 317–367.

Dahan, D., Magnuson, J. S., Tanenhaus,M. K., & Hogan, E. M. (2001). Tracking the time course of subcategorical mismatches: Evidence for lexical competition. *Language and Cognitive Processes, 16,* 507–534.

Dahan, D., Tanenhaus, M. K., & Salverda, A. P. (2007). *The influence of visual processing on phonetically driven saccades in the "visual world" paradigm*. Amsterdam, Netherlands: Elsevier.

Dahan, D., & Gaskell, M. G. (2007). The temporal dynamics of ambiguity resolution: Evidence from spoken-word recognition. *Journal of Memory and Language, 57*(4), 483-501.

Davis, M.H., Marslen-Wilson, W.D., & Gaskell, M. G. (2002). Leading up the lexical garden path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 218-244.

Frauenfelder, U.H., & Peeters, G. (1990). Lexical segmentation in TRACE: An exercisein simulation. In G.T.M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives*. Cambridge, MA: MIT Press.

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance, 6*(1), 110-125.

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation, 12*(4), 395-427.

Gitelman, D. R. (2002). ILAB: A program for postexperimental eye movement analysis. *Behavior Research Methods, Instruments & Computers.Special Issue: Eye movement research methods, 34*(4), 605-612.

Goldinger, S. D. (1996). Auditory lexical decision. *Language and Cognitive Processes, 11*(6), 559-567.

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & psychophysics, 28*(4), 267-283.

Ide, N., Suderman, K. (2004). The American National Corpus First Release. *Proceedings of the Fourth Language Resources and Evaluation Conference (LREC)*, Lisbon, 1681-84.

Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language, 33*(5), 630-645.

Lehiste, I. (1972). The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America, 51,* 2018–2024.

Magnuson, J. S., Dixon, J. A., Tanenhaus, M. K., & Aslin, R. N. (2007). The dynamics of lexical competition during spoken word recognition. *Cognitive Science: A Multidisciplinary Journal, 31*(1), 1-24.

Magnuson, J. S., & Strauss, T. J. (2007). Do spoken words tell us how long they will be? *Twentieth Annual CUNY Conference on Human Sentence Processing.* San Diego

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition. Special Issue: Spoken word recognition, 25*(1-2), 71-102.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive psychology, 10*(1), 29-63.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive psychology, 18*(1), 1-86.

Pitt, M. A., & Samuel, A. G. (2006). Word length and lexical activation: Longer is better. *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1120-1135.

Salverda, A. P., Dahan, D., & McQueen, J. M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition, 90*(1), 51-89.

Samuel, A. G. (1996). Does lexical information influence the perceptual restoration of phonemes? *Journal of Experimental Psychology: General, 125*(1), 28-51.

Singer, J.D., & Willett, J.B. (2003). *Applied longitudinal analysis: Modeling change and event occurrence.* New York: Oxford University Press.

Storkel, H. L. (2004). Methods for minimizing the confounding effects of word length in the analysis of phonotactic probability and neighborhood density. *Journal of Speech, Language, and Hearing Research*, 47(6), 1454-1468.

Vitevitch, M.S. & Luce, P.A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, 40,374-408.

Vitevitch, M. S. (2002). Influence of onset density on spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 270–278.

Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science, 167*(3917), 392-393.