Title
Investigating the Locus of the Word Frequency Effect in Spoken Word Recognition

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 34(34)

ISSN
1069-7977

Authors
Siew, Cynthia
Yap, Melvin
Goh, Winston

Publication Date
2012

Peer reviewed

Investigating the Locus of the Word Frequency Effect in Spoken Word Recognition

Cynthia S. Q. Siew (cynsiewsq@gmail.com)
Melvin J. Yap (melvin@nus.edu.sg)
Winston D. Goh (psygohw@nus.edu.sg)
Department of Psychology, National University of Singapore, 9 Arts Link,
Singapore 117570, Singapore

Abstract

The present study aims to isolate the locus of the frequency effect within the spoken word recognition architecture. By applying the additive factors logic (Sternberg, 1969) to an auditory lexical decision task where both word frequency and stimulus quality were factorially manipulated, the reaction time data can be analyzed to study processing stages along the time course of spoken word recognition, and determine if frequency has an early or late locus. A significant underadditive interaction of frequency and stimulus quality was obtained. Surprisingly, the typically robust frequency effect was not reliable for words of low stimulus quality. This finding suggests that word frequency influences a relatively late stage in the spoken word recognition process. Implications for extant models of spoken word recognition are discussed.

Keywords: Spoken word recognition; word frequency effects; stimulus quality effects; additive factors logic; auditory lexical decision.

Introduction

Determining whether word frequency has an early or late locus has profound theoretical implications for models of spoken word recognition (SWR). While it is well established that frequently occurring words are recognized faster than less frequently occurring words (Goldinger, 1996), what is less obvious is where the *locus* of the frequency effect lies within the word recognition process. Specifically, does frequency influence word recognition at an early stage, as the speech signal begins to unfold, or does frequency influence word recognition at a later stage in the form of a bias? Models of SWR can easily account for the frequency effect, but they do not necessarily agree on the locus of the frequency effect due to varying assumptions and architectures. Hence one way to test the validity of these models is to isolate the locus of the frequency effect.

Several researchers have investigated this issue by employing a variety of experimental techniques and methodologies. Generally, studies which used traditional behavioral experiments (e.g., lexical decision and word identification) have demonstrated that word frequency has a late locus that occurs after lexical processes are complete (Broadbent, 1967; Connine, Mullennix, Shernoff & Yelen, 1990); Luce & Pisoni, 1998). On the other hand, recent studies employing eyetracking technology (e.g., Dahan, Magnuson & Tanenhaus, 2001) and novel behavioral applications of the parallel refractory period paradigm (Cleland, Gaskell, Quinlan & Tamminen, 2006) concluded that word frequency exerts early and facilitatory effects on word recognition. With overwhelming evidence supporting both sides of the debate, the question as to whether word frequency affects spoken word recognition at an early or late stage continues.

The present study aims to isolate the locus of the frequency effect in spoken word recognition by making use of the additive factors logic to investigate this particular research question. The additive factors logic (Sternberg, 1969) is widely used by cognitive psychologists to interpret RT data in factorial experiments and study the stages of processing in a number of research topics (e.g., Stanovich & Pachella, 1977), as the logic can be easily applied in the study a wide array of research topics, including psycholinguistics (e.g., Yap & Balota, 2007).

Additive Factors Logic

According to the additive factors logic (Sternberg, 1969), when two factors affect theoretically determined independent stages in the information processing stream, it should result in *additivity* in mean RTs (i.e., two main effects for each factor, but no interaction). This is represented in the top part of Figure 1, where Factor A affects processing at only Stage 1 and Factor B affects processing at only Stage 2. On the other hand, if the two factors affect the same stage in the information processing stream, this results in a statistical interaction (more precisely, an *overadditive* interaction where the effect of one factor is larger on the “slower” level of the second factor). This is depicted in the bottom part of Figure 1, where both Factor A and Factor B affect processing at a common Stage X.

How can the incorporation of an additional variable, stimulus quality within the auditory lexical decision task, allow us to isolate the locus of the word frequency effect? How can the additive factors logic be used to help us make specific hypotheses about the pattern of results for RT data? In contrast to the lack of consensus with regards to the locus of the frequency effect, few would question the notion that stimulus quality has an early locus of influence in the word recognition process. In fact, a major assumption of most SWR models (e.g., TRACE) involves a process which converts physical, acoustic input into phonemic information (McClelland & Elman, 1986). This necessarily implies that degraded input must be normalized at a relatively early point in the word recognition process.

Hence, if we assume that stimulus quality affects an early stage in the word recognition process, then Factor A corresponds to stimulus quality and Factor B corresponds to frequency, as shown in Figure 2. Hence, additivity (i.e., main effects of frequency and stimulus quality, but no interaction) indicates that stimulus quality and frequency have independent loci of influence, and this further implies that frequency affects a later stage (one that occurs after stimulus quality; as shown in the upper section of Figure 2). An overadditive interaction where the frequency effect is greater for words of low stimulus quality as compared to high stimulus quality would indicate that these two variables influence at least one stage in common, and it follows that frequency has an early locus of influence (as shown in the bottom of Figure 2).

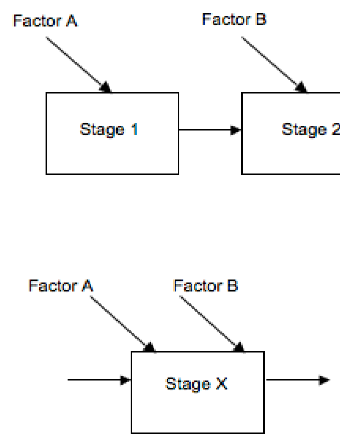

Figure 1. Sternberg's (1969) Additive Factors Logic

In fact, within the area of visual word recognition, some studies have employed the additive factors logic to investigate the joint effects of stimulus quality and frequency in lexical decision (Borowsky & Besner, 1993; Yap & Balota, 2007). Researchers have consistently found that frequency and stimulus quality have *additive* effects in visual lexical decision. This finding is best accommodated within a two-stage model where stimulus quality influences an early stage and frequency influences the second stage (Borowsky & Besner, 1993), which implies that processing at earlier stages is not necessarily frequency-sensitive.

It is also interesting to note that, despite extensive research involving perceptual identification and auditory lexical decision paradigms, researchers almost universally study the effect of stimulus quality on identification accuracy, but not on response latencies. The study of the effects of stimulus quality on spoken word recognition has been largely limited to perceptual and tone identification experiments (Broadbent, 1967; Hawkins & Stevens, 1950; Luce & Pisoni, 1998; Savin, 1963). To our knowledge, stimulus quality has never been directly manipulated as an independent variable in auditory lexical decision**,** and the joint effects of frequency and stimulus quality have not been previously studied in any auditory word recognition task. Hence, another objective of the present study is to address these gaps in the literature.

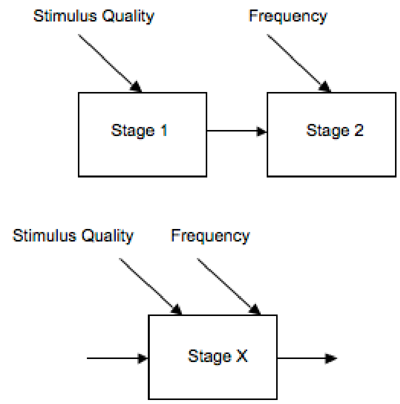

Figure 2. Hypothetical diagrammatic representation of the loci of stimulus quality and frequency effects

Method

Participants

Eighty National University of Singapore undergraduates participated in this study for course credit. Participants' first language was English, and they had no previous reported history of speech or hearing disorders.

Design

A 2 (word frequency: high, low) × 2 (stimulus quality: clear, degraded) mixed-design was used. The within-participants independent variable was word frequency and the between-participants independent variable was stimulus quality. Stimulus quality was manipulated as a between-participants variable to minimize possible carry-over effects that may occur in a fully within-participants design (Poulton, 1982). The dependent variables were reaction time (RT) and accuracy.

Stimuli

Table 1 shows a summary of the descriptive statistics for word and nonword stimuli. 58 high frequency and 58 low frequency English words were selected as stimuli. Using LogFreqHal values generated from the English Lexicon Project (ELP; Balota *et al.*, 2007), the difference between the high and low frequency conditions was reliable, $F(1,114) = 329.72$, $MSe = 273.29$, $p < .001$. High and low frequency words were also matched on number of phonemes, number of syllables, phonological neighborhood density, familiarity rating, uniqueness point, and word duration. A one-way between-items ANOVA showed that for all lexical characteristics, Fs < 1.

116 nonwords were constructed and matched with words on number of phonemes, number of syllables, duration and baseword phonological neighborhood density (an estimation of the nonword density based on the neighborhood density of its closest sounding word). The difference between words and nonwords on each of those variables was not significant, all Fs < 1. All stimuli were spoken by a linguistically trained female speaker and digitally recorded in 16-bit mono, 44.1kHz, .wav format.

Table 1: Descriptive statistics of word and nonword stimuli.

Lexical Characteristics	Word Stimuli				Nonword Stimuli	
	High Frequency		Low Frequency			
	M	*SD*	*M*	*SD*	*M*	*SD*
Number of phonemes	3.95	0.98	4.12	1.04	4.03	1.01
Number of syllables	1.48	0.60	1.43	0.53	1.46	0.57
Phonological neighbourhood density (Baseword phonological density for nonwords)	6.98	6.80	7.31	7.69	7.15	7.23
Duration	582	91	590	81	586	86
Familiarity rating	6.99	0.07	6.98	0.06		
LogFreqHal	11.37	0.75	8.30	1.05		
Uniqueness Point	4.33	0.87	4.26	0.79		

Degrading auditory stimuli. White noise was used to degrade the spoken stimuli, in accordance with past perceptual identification experiments (Broadbent, 1967; Savin, 1963). All degraded trials were presented at SNR +10dB, with white noise at 70dB and target stimuli at 80dB.

Phonemic distributions. Various studies have shown that white noise has differential masking effects on different phonemes (Horii, House & Hughes, 1970; Pisoni, 1996). Following Chan and Vitevitch (2007), chi-square analyses were conducted on the onset consonants, vowels and fricatives of all word stimuli to ensure that no single phoneme was overrepresented among them. The phonetic transcriptions of each word were obtained from the ELP, and subsequent chi-square analyses were not significant.

Procedure

Participants were tested on individual PCs in groups no larger than five. Forty subjects were assigned to the clear condition (without noise) and forty participants were assigned to the degraded condition (with noise). Stimuli were binaurally played through BeyerDynamic DT150 headphones, and E-prime 1.2 software and the PST serial response box (Schneider, Eschman & Zuccolotto, 2002) were used for stimuli presentation and data collection. Participants were instructed to listen to the stimuli carefully and decide, as quickly and accurately as possible, whether the token was a word or a nonword, using the right- and left-most buttons respectively on the response box. Prior to the actual experiment, participants were given 20 practice trials which were not included in the subsequent analyses. For degraded trials, white noise was played for 100ms before the stimulus was presented and continued until 100ms after stimulus offset. Once a response was made, 500ms elapsed before the initiation of another trial. Latencies were measured from the onset of the stimulus until button press. There were a total of 232 experimental trials and participants were allowed a short break after every 58 trials.

Results

For the reaction time data, only correct word trials with RTs more than 200ms and less than 3000ms were included in the analyses. Trials with RTs less than 200ms were excluded, and trials with RTs more than 3000ms were substituted with 3000ms and included in the analysis. This reduces the amount of data excluded and ensures that extreme scores are preserved while reducing their impact (e.g., Marian, Blumenfeld & Boukrina, 2008). Following which, the overall mean and *SD* of each participant's RT was calculated and trials with latencies that were 3 *SD*s above or below each participant's mean RT were removed. These trimming criteria resulted in the removal of 10.2% of all word trials.

The average RTs and Accuracy across the 4 conditions are summarized in Table 2. A two-way mixed-design ANOVA was conducted on the RT and accuracy data, by participants and items.

Table 2: Mean RTs (ms) and accuracy (proportion)

	RT		Accuracy	
	M	*SD*	*M*	*SD*
Clear trials				
High frequency	890	64	0.96	0.03
Low frequency	908	65	0.92	0.04
Frequency effect	18		0.04	
Degraded trials				
High frequency	997	130	0.90	0.05
Low frequency	1002	129	0.87	0.07
Frequency effect	5		0.03	
Interaction	13		0.01	
Nonwords	1237	260	0.76	0.15

Reaction Time

A reliable main effect of stimulus quality, $F_p(1,78) = 19.56$, $MSe = 400656.27$, $p < .001$; $F_i(1,114) = 263.64$, $MSe = 617419.13$, $p < .001$, was found for both participant and item analyses. Across high and low frequency words, participants were slower at recognizing words presented with noise (*M* = 999, *SD* = 129) than for words presented in the clear (*M* = 899, *SD* = 64). The main effect of frequency was significant by participants, $F_p(1,78) = 13.91$, $MSe = 5029.26$, $p < .001$, but not by items, $F_i < 1$. Across both conditions of stimulus quality, response latencies for high frequency words (*M* = 944, *SD* = 115) were significantly faster than response latencies for low frequency words (*M* = 955, *SD* = 112).

The Frequency × Stimulus Quality interaction was significant by participants, $F_p(1,78) = 4.20$, *MSe* = 1517.33, p < .05, but not by item analyses, $F_i(1,114) = 2.14$, *MSe* = 5019.49, *ns*.

Tests of the simple main effect of frequency was significant in the clear condition, *F*(1,78) = 16.69, *MSe* = 6035.73, *p* < .001, but not in the degraded condition, *F*(1,78) = 1.41, *MSe* = 510.86, *ns*. Participants recognized clear high frequency words (*M* = 890, *SD* = 64) more quickly than clear low frequency words (*M* = 908, *SD* = 65). However, for degraded words, participants did not differ on their response latencies for high (*M* = 997, *SD* = 130) and low frequency words (*M* = 1002, *SD* = 129). There was an 18 ms frequency effect at the clear condition, but this was abolished at the degraded level. Tests of the simple main effect of stimulus quality was significant for both high frequency, *F*(1,78) = 21.57, *MSe* = 225743.00, *p* < .001, and low frequency conditions, *F*(1,78) = 16.99, *MSe* = 176430.60, *p* < .001. Among high frequency words, participants were slower to recognize degraded words (*M* = 997, *SD* = 130) than clear words (*M* = 890, *SD* = 64). Among low frequency words, participants were also slower to recognize degraded words (*M* = 1002, *SD* = 129) than clear words (*M* = 907, *SD* = 65).

Accuracy

A reliable main effect of stimulus quality was also found for both participants and items, $F_p(1,78) = 37.52$, *MSe* = .14, *p* < .001; $F_i(1,114) = 34.68$, *MSe* = .20, *p* < .001. Participants were more accurate at recognizing high and low frequency words presented in the clear (*M* = 0.94, *SD* = 0.04) than in the degraded condition (*M* = 0.88, *SD* = 0.06). The frequency effect was reliable by participants, $F_p(1,78) = 30.38$, *MSe* = .04, *p* < .001, and by items, $F_i(1,114) = 4.09$, *MSe* = .06, *p* < .05. Across both levels of stimulus quality, accuracy rates for high frequency words were higher (*M* = 0.93, *SD* = 0.05) than for low frequency words (*M* = 0.90, *SD* = 0.06). No interaction was observed for frequency and stimulus quality in both analyses by participants and by items, $F_p(1,78) = 2.18$, *MSe* = .003, *ns*; $F_i < 1$.

Discussion

In the present study, the joint effects of stimulus quality and word frequency are characterized by an underadditive interaction, as the frequency effect for words of high stimulus quality was reliable but not for words of low stimulus quality. This finding may be considered counterintuitive because additive factors logic does not *a priori* predict underadditivity between two factors. According to additive factors, a statistical interaction is indicative of both variables influencing at least one stage in common in the processing architecture. This interpretation was based on an *overadditive* interaction (Sternberg, 1969), where the effect of one factor is larger at the "slower" level of the second factor. However, the interaction observed here was an *underadditive* one, where the effect of one factor is smaller, instead of larger, at the "slower" level of the second factor.

Consider Figure 3 below, where stimulus quality influences Stage 1 and frequency influences Stage 2. For clear words, word recognition proceeds from Stage 1 to Stage 2. For degraded words, degradation could have slowed processing to the extent that the optional Stage 2 is not initiated. Note that if we assume Stage 2 to be optional and is presumably not necessary for word recognition, then word recognition can still take place without engaging this frequency-sensitive stage. This is consistent with the hypothesis that stimulus quality and frequency influence *separate* processing stages. According to this interpretation, word frequency has a *late* locus of influence that occurs after that of stimulus quality.

It is interesting to note that the underadditive interaction between frequency and stimulus quality parallels the findings of previous studies which have studied the joint effects of frequency and neighborhood density in auditory lexical decision (Goh, Suarez, Yap, & Tan, 2009; Luce & Pisoni, 1998; Metsala, 1997). These studies found that frequency and neighborhood density interact underadditively, and frequency effects are attenuated for words belonging to dense neighborhoods. This appears to correspond with our present finding that the frequency effect was not reliable for degraded words, as both results indicate that frequency effects are *smaller* when word processing is *slowed* down, either via degradation or neighborhood density effects.

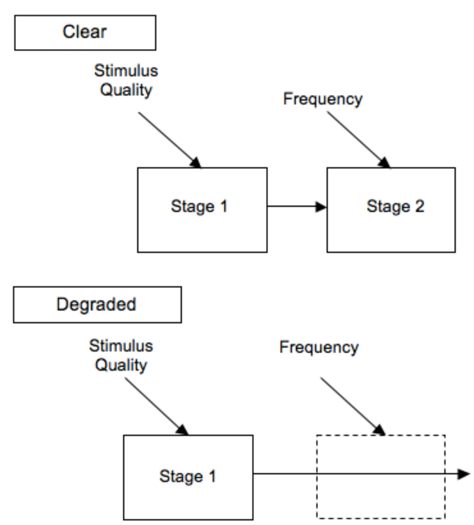

Figure 3. Diagrammatic representation of the underadditive effects of frequency and stimulus in a two-stage model

To some extent, degraded words are analogous to words belonging to dense neighborhoods. Words in dense neighborhoods have several phonological neighbors, which are defined as words differing from the target word on at least one phoneme in any position (Yates, 2005). The acoustic-phonetic patterns of these words are likely to be more confusable because there are several words sharing similar patterns (Luce & Large, 2001). Hence, words with several neighbors tend to activate many more word units than words with fewer neighbors. This results in more competition among word units which inhibits word recognition performance for high density words (Luce & Pisoni, 1998).

Therefore, it is possible that introducing degradation to word stimuli similarly increases the level of competition among word units. Due to increased acoustic ambiguity, a degraded stimulus can be potentially matched to a large number of words in the lexicon and this leads to a large number of potential word candidates being activated and subsequently competing for recognition. In general, it appears that the ambiguity of the acoustic input (due to either exogenous noise or the perception of the acoustic input as possibly corresponding to several word units) ultimately leads to an increase in competition among activated word units, slowing processing to the point that any biasing effects of word frequency are not observed.

Implications for Models of SWR

The present study is of theoretical importance because the results can impose additional constraints on speech recognition models. In this section the implications of the underadditive interaction for extant models of SWR are briefly reviewed.

To account for the finding that frequency effects are attenuated in certain tasks (e.g., Connine *et al.* 1990), NAM posits that frequency effects are non-obligatory and that it is possible for word recognition to occur *without* involving the later, frequency-sensitive stage (Luce & Pisoni, 1998). NAM conceptualizes the frequency effect as a decision bias that occurs later in the word recognition process, and it appears that this bias can be "turned off" depending on the task demands and conditions. Therefore, NAM is able to accommodate the present finding because in this model word recognition can still occur with limited or no processing at Stage 2 (see Figure 3), and this explains why a reliable frequency effect was not obtained for degraded words.

Other models of SWR are unable to accommodate the present finding as easily. In the TRACE model of speech perception, (McClelland & Elman, 1986), since frequency and stimulus quality both influence an architecture that allows bidirectional flow of information between processing levels, it should predict an overadditive interaction as this is analogous to two variables influencing a common stage (the influences of word frequency or stimulus quality are not independent of each other). In order to accommodate the underadditive interaction, we speculate that the model needs to allow for a flexible word processing system that can reduce the influence of frequency when the acoustic input is compromised such that bottom-up flow of perceptual evidence is considerably slowed down.

The results are also inconsistent with the predictions of another major model of SWR - Shortlist B. According to this model, optimal listeners rely more on prior probabilities to compute conditional probabilities (using Bayes Theorem) when ambiguity of the speech input is high (Norris & McQueen, 2008). Since prior probabilities of words are approximated to word frequency and adding white noise to spoken stimuli increases the perceptual uncertainty of the speech input, the model predicts that the frequency effect should be larger for degraded words compared to clear words (Norris & McQueen, 2008). However, in this study, the frequency effect was abolished for degraded words, which seems to imply that listeners actually rely *less* on prior probabilities under increased perceptual uncertainty.

In summary, to account for the non-reliable frequency effect in the degraded condition, we proposed that this was due to degradation inducing a high level of competition among word candidates, such that the frequency-sensitive stage is not invoked over the course of spoken word recognition. Therefore, the finding of an underadditive, rather than overadditive, interaction between stimulus quality and frequency can be accommodated by a two-stage model where the second, frequency-sensitive stage is not mandatory for word recognition. This further suggests that these variables influence *separate* stages in the word recognition process, and by extension, that word frequency has a late locus of influence occurring after that of stimulus quality.

Notes

As suggested by a reviewer, we conducted mix-effects modelling on our data using R (R Development Core Team, 2011). A linear mixed effects model was fitted to the RT data from the experiment, using the lme4 package (Bates *et al.*, 2012); *p*-values for fixed effects were computed using the languageR package (Baayen, 2012). The main effects of stimulus quality and frequency, and the interaction between the two factors were treated as fixed effects, while participants and items were treated as random variables. Our results revealed a significant main effect of stimulus quality ($p < .001$) and no effect of frequency. These were qualified by a marginally significant stimulus quality by frequency interaction, $p = .070$.

Acknowledgements

This work was supported by Research Support Scheme C-581-000-222-091 to WDG. We thank Eileen Soh for data collection assistance.

References

Baayen, R. H. (2012). languageR: Data Sets and Functions with "Analyzing Linguistic Data: A Practical

Introduction to Statistics," R Package Version 1.4. Vienna: R Foundation for Statistical Computing.

Balota, D.A., Yap, M.J., Cortese, M.J., Hutchison, K.A., Kessler, B., Loftis, B., Neely, J.H., Nelson, D.L., Simpson, G.B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39*, 445-459.

Bates, D. M., Maechler, M., & Dai, B. (2012). lme4: Linear Mixed-Effect Models Using S4 Classes, R Package Version 0.999375-42. Vienna: R Foundation for Statistical Computing.

Borowsky, R., & Besner, D. (1993). Visual word recognition: A multistage activation model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 813-840.

Broadbent, D. E. (1967). Word-frequency effect and response bias. *Psychological Review*, *74*, 1-15.

Chan, K. Y., & Vitevitch, M. S. (2007). The influence of the phonological neighborhood clustering coefficient on spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 1934-1949.

Cleland, A. A., Gaskell, M. G., Quinlan, P. T., & Tamminen, J. (2006). Frequency effects in spoken and visual word recognition: Evidence from dual-task methodologies. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 104-119.

Connine, C. M., Mullennix, J., Shernoff, E., & Yelen, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*,, 1084-1096.

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.

Goh, W. D., Suarez, L., Yap, M. J., & Tan, S. H. (2009). Distributional analyses in auditory lexical decision: Neighborhood density and word-frequency effects. *Psychonomic Bulletin & Review*, *16*, 882-887.

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

Hawkins, J. E., & Stevens, S. S. (1950). The masking of pure tones and of speech by white noise. *Journal of the Acoustical Society of America*, *22*, 6-13.

Horii, Y., House, A. S., & Hughes, G. W. (1970). A masking noise with speech-envelope characteristics for studying intelligibility. *Journal of the Acoustical Society of America*, *49*, 1849-1856.

Luce, P. A., & Large, N. R. (2001). Phonotactics, density, and entropy in spoken word recognition. *Language and Cognitive Processes, 16*, 565-581.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing Spoken Words: The Neighborhood Activation Model. *Ear & Hearing*, *19*, 1-36.

Marian, V., Blumenfeld, H. K., & Boukrina, O. V. (2008). Sensitivity to phonological similarity within and across languages. *Journal of Psycholinguistic Research, 37,* 141-170.

McClelland, J. L., & Elman, J. L. (1986). The TRACE Model of Speech Perception. *Cognitive Psychology*, *18*, 1-86.

Metsala, J. L. (1997). An examination of word frequency and neighborhood density in the development of spoken-word recognition. *Memory & Cognition*, *25*, 47-56.

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, *115*, 357-395.

Pisoni, D. B. (1996). Word identification in noise. *Language and Cognitive Processes*, *11*, 681-687.

Poulton, E. C. (1982). Influential companions: Effects of one strategy on another in the within-subjects designs of cognitive psychology. *Psychological Bulletin*, *91*, 673-690.

R Development Core Team. (2011). R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing.

Savin, H. B. (1963). Word-frequency effect and errors in the perception of speech. *Journal of the Acoustical Society of America*, *35*, 200-206.

Schneider, W., Eschman, A., & Zuccolotto, A. (2001). *E-prime user's guide*. Pittsburgh: Psychology Software Tools, Inc.

Stanovich, K. E., & Pachella, R. G. (1977). Encoding, stimulus-response compatibility, and stages of processing. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 411-421.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, *30*, 276-315.

Yap, M. J., & Balota, D. A. (2007). Additive and interactive effects on response time distributions in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 274-296.

Yates, M. (2005). Phonological neighbors speed visual word processing: Evidence from multiple tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1385-1397.