Modelling Graded Semantic Effects in Lexical Decision

Ya-Ning Chang (ya-ning.chang@manchester.ac.uk)
Neuroscience and Aphasia Research Unit (NARU), University of Manchester,
Brunswick Street, Manchester, M13 9PL, UK

Matthew Lambon Ralph (matt.lambon-ralph@manchester.ac.uk)
Neuroscience and Aphasia Research Unit (NARU), University of Manchester,
Brunswick Street, Manchester, M13 9PL, UK

Steve Furber (steve.furber@manchester.ac.uk)
Advanced Processor Technologies Group, University of Manchester, Oxford Road, Manchester, M13 9PL, UK

Stephen Welbourne (stephen.welbourne@manchester.ac.uk)
Neuroscience and Aphasia Research Unit (NARU), University of Manchester,
Brunswick Street, Manchester, M13 9PL, UK

Abstract

Recent studies have shown that the involvement of semantic information in visual lexical decision depends on the nature of nonword foils with semantic effects increased as nonwords become more word-like (Evans, Lambon Ralph &Woollams, 2012). Given that most models of lexical decision focus on orthographic information (Coltheart, Rastle, Perry, Langdon & Ziegler, 2001; Grainger & Jacobs, 1996; Seidenberg & McClelland, 1989), the role of semantics and its interactions with vision, orthography, and phonology has been overlooked. We developed a recurrent connectionist model of single word reading including visual, orthographic, phonological, and semantic processing. The model differentiated words from nonwords by integrating measures of polarity across four key processing layers. The contribution of semantics depended on the type of nonword foils. The model was more reliant on semantic information when the nonword foils were pseudowords and pseudohomophones rather than consonant strings. The results support the view that semantic involvement in lexical decision is graded by the difficulty of the decision task.

Keywords: semantic effects; lexical decision; reading; computational modelling; visual word recognition.

Introduction

Lexical decision (LD) has been widely used to study the cognitive processes involved in visual word recognition. Subjects are asked to judge whether a letter string is a word or not. Measures of accuracy and response time are thought to reflect the differences in lexical-semantic processing of words and nonwords. There seems to be consistent evidence that vision, orthography and phonology play roles in visual lexical decision (Coltheart, Davelaar, Jonasson, & Besner, 1977; Grainger & Jacobs, 1996; Meyer, Schvanev, & Ruddy, 1974), however the extent of the involvement of semantics in lexical decision remains debateable (James, 1975; Joordens & Becker, 1997; Lupker & Pexman, 2010). James (1975) showed a reliable concreteness effect during lexical decision when using pseudoword and pseudohomophone foils, while the effect disappeared when testing with consonant strings. He suggested subjects might be able to exploit semantic information to support efficient LD. Although some subsequent studies have found reliable semantic influences on lexical decision under different foil conditions (Joordens & Becker, 1997), others have failed to find such effect (Lupker & Pexman, 2010). Evans, Lambon Ralph and Woollams (2012) demonstrated that semantic involvement in lexical decision was graded by the difficulty of the decision task as indexed by the word-likeness of the foil. There were stronger semantic effects with pseudohomophones than with pseudowords, and the effects were stronger with pseudowords than with consonant strings. Apart from the behavioural data, there is also evidence of semantic involvement in lexical decision from neuroimaging studies. Woollams, Silani, Okada, Patterson and Price (2011) revealed that left anterior temporal activation, increased for atypical relative to typical strings when lexical decisions were made more difficult in the context of pseudohomophone foils. The left anterior temporal lobe has been considered as a region for combining various types of sensory and motor information to form amodal semantic representations (Patterson, Nestor, & Rogers, 2007). The orthographic typicality effect in the left anterior temporal lobe has also been found in a previous electrophysiological (EEG) study. In a speeded lexical decision task, atypical words were found to elicit stronger source currents than did typical words at around 160 msec in the left anterior temporal lobe (Hauk, Patterson, Woollams, et al., 2006). These effects are consistent with what has been observed in the neuropsychological studies of patients with semantic dementia (SD), who have asymmetrically bilateral atrophy degeneration of the anterior temporal lobes. These patients show a progressive degeneration of semantic knowledge (Hodges, Patterson, Oxbury, & Funnell, 1992). When patients are asked to perform two-alternative forced-choice visual lexical decision, they can correctly choose orthographically typical words from the relatively atypical nonwords but have difficulty in the reverse condition (Rogers, Lambon Ralph, Hodges, & Patterson, 2004). Taken together, this evidence supports the view that semantic processing is involved in lexical decision in particular when the words are orthographically atypical and the foils are pseudohomophones.

Models Based on Localist Views

In the literature, several theories of visual word recognition have been proposed to explain the underlying mechanisms of lexical decision (Coltheart, et al., 1977; Coltheart, et al., 2001; Grainger & Jacobs, 1996; Plaut, 1997; Seidenberg & McClelland, 1989). Some researchers argue that lexical decision relies upon the orthographic lexicon (Coltheart, et al., 1977). If there is a match, subjects would give a positive response, otherwise, the negative response is made. On this view, the locus of lexical decision is based on activation within the orthographic lexicon. The involvement of phonology is a relatively late process after the mental lexicon search while the semantic system is generally not involved in the recognition processes unless the discrimination becomes extremely difficult (Coltheart, et al., 2001). This orthographically based approach is shared with Grainger and Jacobs (1996), who developed a computational model of lexical decision. In their multiple read-out model (MROM), a word response could be made either when the particular word unit activation reached a local criterion, M, or the overall activity in the word layer reached a global criterion, Σ, before the temporal deadline as T. The RT was based on the earliest moment where either of criteria was met. If neither of the activation criteria was met, a nonword response was given and the RT was the value of the deadline criterion. Grainger and Jacobs (1996) assumed that the M criterion should be fixed as a normal recognition level and was set corresponding to individual word units. While the global criterion Σ and the temporal deadline T would vary according to the lexical frequency status of the stimulus. The higher probability the stimulus was a word, the lower global criterion and the longer temporal deadline were used. By this, the MROM model was able to simulate several standard effects seen in lexical decision including the frequency effects, the orthographic neighbourhood size effects, and their interactions (Grainger and Jacobs, 1996). Other models of visual word recognition such as the dual-route cascaded (DRC) model (Coltheart, et al., 2001) and the connectionist dual process (CDP+) model (Perry, Ziegler, & Zorzi, 2007) share similar decision mechanisms to the MROM model.

Models Based on Distributed Views

An alternative theory of visual word recognition argues that there is no mental lexicon for the store of word knowledge in the recognition system (Dilkina, McClelland, & Plaut, 2010; Plaut, 1997; Seidenberg & McClelland, 1989). On this view, the decision can be made on the basis of the differential activations elicited by familiar words and unfamiliar nonwords. When presenting a word, strong activations are expected because the mappings between the visual or orthographic representation of the word and its phonological and semantic representations have been learned. Conversely, relatively weaker activations would be expected for a nonword representation as it is a novel stimulus. One important model of lexical decision was developed by Plaut in 1997, who proposed that the measure of how strongly units were activated, called stress or polarity, could be used as a basis for making lexical decisions. He built a feedforward model which consisted of orthographic, phonological and semantic components and demonstrated that words tended to produce higher stress than nonwords at the semantic layer. With the proper decision criteria, over 95 percent of words in the training corpus could be discriminated from nonwords. In addition, the network tended to produce higher semantic stress for pseudohomophones than for pseudowords in line with the behavioural data.

Accumulated Information for Lexical Decision

There are also other models which have emphasised the use of accumulated information within the system for making decisions. One of these is the diffusion model, developed by Ratcliff, Gomez and Mckoon (2004). The central idea of the diffusion model was that the speed (drift rate) at which information was accumulated over time was affected by the lexical status of the stimuli. They hypothesized that the drift rate had a positive correlation with a measure of how word-like a stimulus was. In their model, the decision was then made when a random walk process driven by the drift rate reached either a word criterion or nonword criterion. Another model is the Bayesian reader model developed by Norris (2009). The basic premise of this model was to assume subjects would consistently compute the probability of the stimulus being a word or a nonword on the basis of its lexical status. In the simulations conducted in Norris (2009), the recognition of a letter string being a word was made on the basis of the sum of the probabilities of all possible letter strings and this value was expected to be 1.0. Therefore, the nonword likelihood could be computed simply by using 1 minus summed probability of letter strings corresponding to words.

In summary, data from behavioural, neuroimaging and patient studies, all point to the involvement of semantic processing in lexical decision. Previous models either postulate an exclusive role for semantics (Dilkina, McClelland, & Plaut, 2010; Plaut, 1997) or no role for semantics (Coltheart et al. 2001; Grainger & Jacobs, 1996; Norris, 2009). Importantly none of the previous models would be able to account for the data from Evans et al. (2012), which indicates that the degree of semantic involvement is flexible and can be modulated by the nature of the nonwords foils. The goal of this paper was to use a novel model of reading to explore to what extent semantics is involved in lexical decision and how it interacts with other processing layers. In addition we aimed to be able to simulate the data from Evans et al. illustrating how changes to the nature of the nonwords foils can bias lexical decision tasks. Based on earlier work (Chang, Furber, & Welbourne, 2012a), we developed a fully implemented recurrent model of visual word recognition. The model included a visual processing stage along with the orthographic, phonological and semantic processing stages. Importantly, the orthographic representations were allowed to learn during the training.

Method

Network Architecture

The architecture of the model is shown in Figure 1. The model had two separate pathways for recognising words from visual input: a phonological pathway and a semantic pathway. The H0 layer was functionally responsible for visual processing while the OH layer was equivalent to the orthographic layer in the triangle model except that the orthographic representations were learned through the course of training rather than being supplied as inputs. This mimics the situation in human development where orthographic representations emerge to support reading acquisition in children. The word recognition process started from the visual input layer and moved progressively to the orthographic layer, and then progressed in separate pathways to the phonological and semantic layers. The phonological component consisted of 61 phonological units which were all connected to a set of 20 clean up units. These clean up units projected back onto the phonological units, forming an attractor. Similarly, the semantic component consisted of 200 semantic units. These units were all connected to another set of 80 clean up units, which projected back onto the semantic units. The context component consisted of 3 units, which were used to provide additional contextual information for discriminating between homophones. The numbers of hidden units for each layer were determined by pilot trials to ensure the model was trainable and that the performance of the model was good on the production, comprehension and reading tasks.

There were also control units for each layer except input and output layers. These acted to flexibly inhibit the activation of the layer they were connected to. The control units were important because they allowed the model learn to manage its own temporal dynamics. In particular they allowed the units at the latter layers to be suppressed until the input to them had had time to ramp up to values that reflected the influence of the visual input to the model.

The training corpus consisted of 2,971 words. The visual representations used here were adapted from those used in Chang et al.'s (2012a) study. The network was trained on 12-point lower case words in Arial font. Each word was positioned with its vowel aligned on a fixed slot of the image. Ten slots were used in all and the size of each slot was 16x16 pixels. The scheme of phonological representations was the same as that used in the Plaut et al.'s (1996) model. The context units were used to differentiate the meanings of homophones, which have same pronunciations but different meanings. For those pronunciations with only one possible word meaning, the context units were all set to zero. For other pronunciations corresponding to more than one word meanings, the context units were all set to 0 for the first meaning; and one of the context units from right to left was set to 1 to represent the second, third and fourth meaning accordingly. The semantic representations were generated using the same scheme as in Chang, Furber, and Welbourne (2012b). The meaning of each word was represented by a 200-dimensional semantic vector. Each vector had 5 active units in the first half of the vector converted from the top positive attributes and 15 active units in the second half of the vector converted from the top negative attributes.

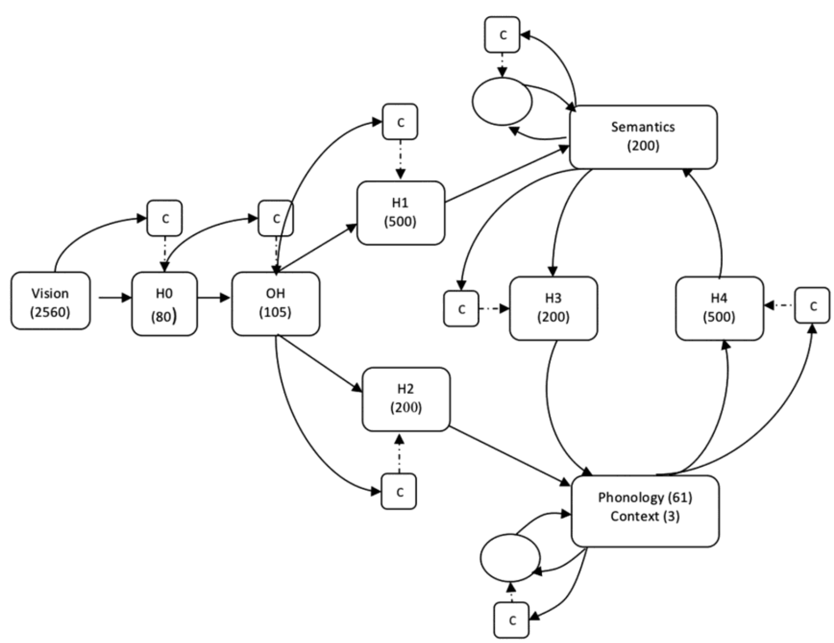

Figure 1. The architecture of the model. The dashed lines indicate inhibitory connections.

Training Procedure

The training was separated into two phases. In phase 1 only the phonology-semantics mappings were trained while in phase 2 the full reading model was trained starting from the trained weights obtained in phase 1. In phase 1, the phonology-semantics model was first subdivided into two parts: the production model learning the mappings from semantics to phonology and context, and the comprehension model learning the mappings from phonology and context to semantics. The production and comprehension model were trained separately. The presentation of each example lasted for 6 intervals of time and each interval of time was divided into 3 ticks. In each presentation, the input pattern of a word was clamped onto the input units for the full 6 intervals of time and the task was to produce the correct target representation. For the last 2 intervals, the activations of output units were compared to their targets. Error score, the difference between the units' outputs and their targets, was used to calculate weight changes. No error was recorded if the output unit's activation and target were within 0.1 of each other. At the end of phase 1 the accuracy rates of the production and comprehension model were 99.97% and 99.43% for the phonological level and semantic level respectively.

In phase 2, the weights obtained from the end of training the phonology-semantics model were embedded and frozen into the full reading model. The weight connections from the visual layer to both phonological and semantic layers were updated through training. There were local control units for each layer except input and output layers. The initial output of each control unit was set to 1. The weight connections from its previous layer to each control unit were free to be updated. The weight connections from each control unit to those units that it was controlling were trainable, but the values were limited to between -4 and 0. The negative boundaries used here were to ensure that the control unit acted to inhibit activation. The model was allowed to update for 30 ticks of time. The visual representation of a word was presented at the input units for all 30 ticks. The task was to produce correct phonological and semantic patterns. For the last 2 intervals, the output units were compared with their corresponding phonological or semantic targets and errors were computed. To encourage more accurate learning, no error was computed when the output unit's activation and target were within 0.001.The model was trained to produce 99.3% correct phonological and 97.4% correct semantic patterns in the word reading task.

Polarity Measures and Decision Criteria

Plaut (1997) proposed that parallel distributed models can perform the lexical decision task based on the measure of polarity, which is whether the units in the model have learned to adopt a binary representation. To capture this phenomenon, Plaut (1997) introduced a formula to compute the index of unit binarization which was termed unit polarity as follows:

$$y = x \cdot log_2(x) + (1 - x) \cdot log_2(1 - x) + 1$$

where x is the unit activation ranging from 0 to 1; $log_2(\cdot)$ is the logarithmic function with the base of 2; y is the polarity measure. When known words are presented, the units tend to become binary, leading to high polarity values. However, when nonwords are presented, the activities of the units tend to be low and closer to 0.5, resulting in generally low polarities. Two criteria were used for the model to make word-nonword decisions: (1) word boundary: the 3 standard deviation line above the average nonword polarity; (2) nonword boundary: the 3 standard deviation line below the average word polarity. The polarity for an item was computed by combining the measures of polarity for that item at the H0 (visual processing), OH (orthographic processing), phonological, and semantic layers. If an item polarity crossed over the word boundary the item was classified as a word. By contrast, if the item polarity crossed over the nonword boundary, the item was determined as a nonword. There were, however, a few item polarities that remained between the two boundaries. In this case, responses were made based on which boundary the polarity was closest to at the last time tick. The response time was the time tick when an item polarity first crossed over either word or nonword boundary. In the situation where neither boundary was crossed the response time was taken as 30 ticks.

Inverse Efficiency

To control for potential differences in speed-accuracy trade-off caused by the arbitrary selection of standard deviation lines, we adopted a measure of inverse efficiency, which is considered to be a corrected reaction time (Roder, Kusmierek, Spence, & Schicke, 2007). Inverse efficiency is a combination of both reaction and accuracy (i.e., dividing reaction time by accuracy). The lower the score, the more efficiently the model performed the task.

Results

Semantic influences on lexical decision

Evans et al. (2012) suggested that the subjects needed to access semantic information in the lexical decision task particularly when words were tested with more word-like nonwords such as pseudowords and pseudohomophones. They showed a graded imageability effect in lexical decision depending on the difficulty of the task. The imageability effect was larger when words were tested against with pseudohomophones than with pseudowords. The effect disappeared in the context of consonant strings. We tested the model to see whether it could produce the similar pattern as seen in Evans et al.'s data. After removing those words which were not in the training exemplars and their matched nonword items, there were 70 words, consisting of 35 high- and 35 low-imageability words. Their matched nonword pairs for the three different foil conditions, consonant string (CS), pseudoword (PW), and pseudohomophone (PH) were also used in the current test. To compare with Evans et al (2012)'s data, the scores of inverse efficiency were normalised by the value obtained from the low imageability pseudohomophone condition. The same procedure was applied to Evans et al.'s (2012) data. The results are shown in Figure 2. It is clear that the simulation results (Figure 2, left) follow the pattern of Evans et al.'s data (Figure 2, right). A 2x3 repeated measures ANOVA was conducted with imageability (High/Low) and foil condition (CS/PW/PH) as within subject factors and the scores of inverse efficiency were used as a dependent variable. There was a reliable main effect of imageability, $F(1, 19)=9.88$, $p<.01$. The main effect of foil condition was also significant, $F(1.31, 24.85)=59.75$, $p<.001$ (with a Greenhouse-Geisser adjustment). Importantly, there was a significant interaction between imageability and foil condition, $F(2, 38)=3.60$, $p<.05$, showing that the size of imageability effect increased along with the word-likeness of the foils. Note that we also ran the statistical tests on the unnormalised scores with the same pattern of results. This is what would be expected based on Evans et al.'s (2012) data. The post-hoc analyses showed that the imageability effect was not significant with consonant strings ($p>.05$) while there were significant imageability effects in the contexts of pseudowords, $F(1, 19)=6.76$, $p<.05$, and pseudohomophones, $F(1, 19)=15.06$, $p<.01$. The results were consistent with the findings in Evans et al.'s (2012) study, suggesting semantic effects vary in lexical decision, depending on the foil type.

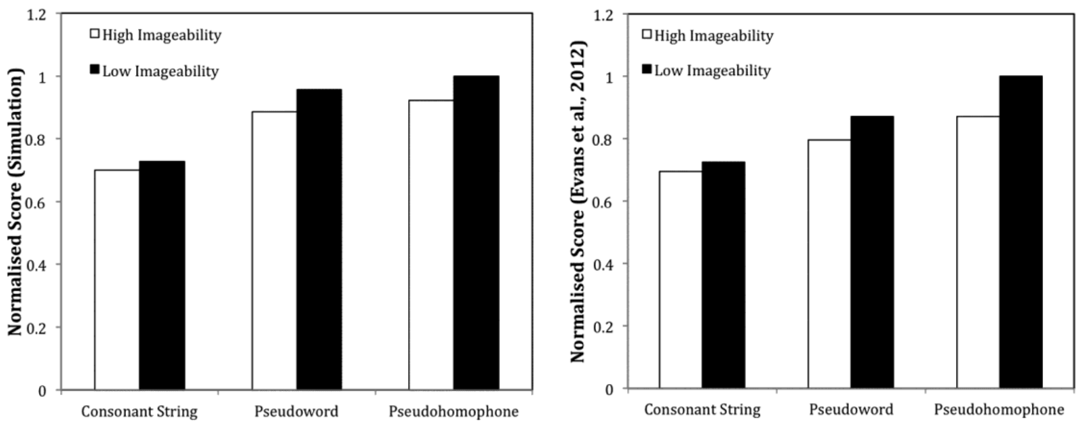

Figure 2. Data are from simulation (Left) and from Evans et al. (2012). Normalised scores were computed by equating two results based on the low imageability pseudohomophone condition.

General Discussion

The primary aim of this paper was to develop a large-scale recurrent reading model containing visual, orthographic, phonological, and semantic processing to support lexical decision tasks. The model was used to explore the involvement of semantics in lexical decision with other processing components implemented in the system. This approach is different to most existing models of lexical processing which have focused on activity within a single processing layer. Based on the measure of polarities at four core processing layers (H0, OH, phonology and semantics), the model was able to perform the lexical decision tasks and account for the graded semantic effects found by Evans et al. (2012), as shown in Figure 2. The magnitude of semantic effects increased as nonwords became more word-like, where the semantic effect was stronger with pseudohomophones than with pseudowords and then with consonant strings. This provides evidence supporting the distributed view of lexical decision which proposes that semantic access is important and automatic in lexical decision (Plaut, 1997). The actual use of semantic information is flexible and is largely dependent on the difficulty of the tasks (Evans, et al., 2012). That is in contrast with the localist view arguing for no or little involvement of semantics in lexical decision (Coltheart, et al., 2001).

There are some existing lexical decision models developed on the basis of the localist view of lexical decision including the MROM model (Grainger & Jacobs, 1996) and the DRC model (Coltheart, et al., 2001) and the CDP+ model (Perry, et al., 2007). These models can simulate several effects in lexical decision and the strategic influences on lexical decision by flexibly adjusting decision criteria. However, their results are almost all based on orthographic processing with little attention to other processing components in particular the semantic system. Thus the questions as to how these models implement the involvement of semantics in lexical decision, which presumably requires some feedback connections from semantics to their orthographic lexicon (Coltheart, et al., 2001) remain unclear. In particular, these localist models would find it difficult to account for the graded changes in the involvement of semantics depending on foil type. In the current model this graded effect emerges naturally as a consequence of increasing task difficulty.

In this paper we have followed Evans et al. (2012) by talking the size of the imageability effect as an index of semantic involvement, but future work could extend this in the model by developing additional metrics to quantify the involvement of semantics including a direct comparison of performance with and without the contribution from the semantic layer.

To summarise, this paper uses a model of human visual word recognition to explore the role of semantics in lexical decision. Crucially, the model was able to account for the graded semantic influences on lexical decision corresponding to the various types of foils, providing evidence for semantic influences on lexical decision.

Acknowledgments

This research was supported by grants under the Cognitive Foresight Initiative (jointly funded by EPSRC, MRC and BBSRC - EP/F03430X/1) and the Neuroscience Research Institute at the University of Manchester.

References

Chang, Y. N., Furber, S., & Welbourne, S. (2012a). "Serial" Effects in Parallel Models of Reading. *Cognitive Psychology, 64*(4), 267-291.

Chang, Y. N., Furber, S., & Welbourne, S. (2012b). Generating Realistic Codes for Use in Neural Network Models. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Confrerence of the Cognitive Science Society* (pp. 198-203). Austin, Tx: Cognitive Science Society.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the Internal Lexicon. In S. Dornic (Ed.), *Attention and Performance Vi*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). Drc: A Dual Route Cascaded Model of Visual Word Recognition and Reading Aloud. *Psychological Review, 108*(1), 204-256.

Dilkina, K., McClelland, J. L., & Plaut, D. C. (2010). Are There Mental Lexicons? The Role of Semantics in Lexical Decision. *Brain Research, 1365*, 66-81.

Evans, G. A. L., Lambon Ralph, M. A., &Woollams, A. M. (2012). What's in a Word? A Parametric Study of Semantic Influences on Visual Word Recognition. *Psychonomic Bulletin & Review*, 19(2), 325-331.

Grainger, J., & Jacobs, A. M. (1996). Orthographic Processing in Visual Word Recognition: A Multiple Read-out Model. *Psychological Review, 103*(3), 518-565.

Hauk, O., Patterson, K., Woollams, A., Watling, L., Pulvermuller, F., & Rogers, T. T. (2006). Q: When Would You Prefer a Sossage to a Sausage? A: At About 100 Msec. Erp Correlates of Orthographic Typicality and Lexicality in Written Word Recognition. *Journal of Cognitive Neuroscience, 18*(5), 818-832.

Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic Dementia - Progressive Fluent Aphasia with Temporal-Lobe Atrophy. *Brain, 115*, 1783-1806.

James, C. T. (1975). Role of Semantic Information in Lexical Decisions. *Journal of Experimental Psychology-Human Perception and Performance, 104*(2), 130-136.

Joordens, S., & Becker, S. (1997). The Long and Short of Semantic Priming Effects in Lexical Decision. *Journal of Experimental Psychology-Learning Memory and Cognition, 23*(5), 1083-1105.

Lupker, S. J., & Pexman, P. M. (2010). Making Things Difficult in Lexical Decision: The Impact of Pseudohomophones and Transposed-Letter Nonwords on Frequency and Semantic Priming Effects. *Journal of Experimental Psychology-Learning Memory and Cognition, 36*(5), 1267-1289.

Meyer, D. E., Schvanev, R. W., & Ruddy, M. G. (1974). Functions of Graphemic and Phonemic Codes in Visual Word-Recognition. *Memory & Cognition, 2*(2), 309-321.

Norris, D. (2009). Putting It All Together: A Unified Account of Word Recognition and Reaction-Time Distributions. *Psychological Review, 116*(1), 207-219.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where Do You Know What You Know? The Representation of Semantic Knowledge in the Human Brain. *Nature Reviews Neuroscience, 8*(12), 976-987.

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested Incremental Modeling in the Development of Computational Theories: The Cdp+ Model of Reading Aloud. *Psychological Review, 114*(2), 273-315.

Plaut, D. C. (1997). Structure and Function in the Lexical System: Insights from Distributed Models of Word Reading and Lexical Decision. *Language and Cognitive Processes, 12*(5-6), 765-805.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding Normal and Impaired Word Reading: Computational Principles in Quasi-Regular Domains. *Psychological Review, 103*(1), 56-115.

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A Diffusion Model Account of the Lexical Decision Task. *Psychological Review, 111*(1), 159-182.

Roder, B., Kusmierek, A., Spence, C., & Schicke, T. (2007). Developmental Vision Determines the Reference Frame for the Multisensory Control of Action. *Proceedings of the National Academy of Sciences of the United States of America, 104*(11), 4753-4758.

Rogers, T. T., Lambon Ralph, M. A., Hodges, J. R., & Patterson, K. (2004). Natural Selection: The Impactof Semantic Impairment on Lexical and Object Decision. *Cognitive Neuropsychology, 21*(2-4), 331-352.

Seidenberg, M. S., & McClelland, J. L. (1989). A Distributed, Developmental Model of Word Recognition and Naming. *Psychological Review, 96*(4), 523-568.

Woollams, A. M., Silani, G., Okada, K., Patterson, K., & Price, C. J. (2011). Word or Word-Like? Dissociating Orthographic Typicality from Lexicality in the Left Occipito-Temporal Cortex. *Journal of Cognitive Neuroscience, 23*(4), 992-1002.