Contemporary Educational Psychology 35 (2010) 126–140

Contents lists available at ScienceDirect

Contemporary Educational Psychology

journal homepage: www.elsevier.com/locate/cedpsych

Word learning: An ERP investigation of word experience effects on recognition and word processing

Michal Balass *, Jessica R. Nelson, Charles A. Perfetti

University of Pittsburgh, Pittsburgh, PA 15260, United States

ARTICLE INFO

Article history:
Available online 13 April 2010

Keywords:
Vocabulary learning
Reading comprehension skill
Word processing
ERP study
N400
P600

ABSTRACT

Adults of varying reading comprehension skill learned a set of previously unknown rare English words (e.g., *gloaming*) in three different learning conditions in which the type of word knowledge was manipulated. The words were presented in one of three conditions: (1) orthography-to-meaning (no phonology); (2) orthography-to-phonology (no meaning); and (3) phonology-to-meaning (no orthography). Following learning, participants made meaning judgments on the learned words, familiar known words, and unpresented (unlearned) rare words while their ERPs were recorded. The behavioral results showed no significant effects of comprehension skill on meaning judgment performance. Contrastingly, the ERP results indicated comprehension skill differences in P600 amplitude; high-skilled readers showed stronger familiarity effects for learned words, whereas less-skilled readers did not distinguish between learned words, familiar words, and unlearned words. Evidence from the P600 and N400 illustrated superior learning of meaning when meaning information was coupled with orthography rather than phonology. These results suggest that the availability of word knowledge (orthography, phonology, and meaning) at learning affects subsequent word identification processes when the words are encountered in a new context.

1. Introduction

Much of the knowledge we have about words and how we process them develops from a lifetime of learning experiences with words. What is known about a word and its meaning and the quality of this knowledge affects how it is processed in isolation, and has even more critical effects on higher level comprehension processes. Theories of reading comprehension, such as the Verbal Efficiency Theory (Perfetti, 1985), and more recently, the Lexical Quality Hypothesis (Perfetti, 2007; Perfetti & Hart, 2001, 2002) propose that difficulties in comprehension partly arise from poor word identification skills and weak representations of a word's form (orthography and phonology) and its meaning. These theories are supported by correlational data that indicate word knowledge to be a critical predictor of reading ability (Dixon, LeFevre, & Twilley, 1988; Ouellette, 2006; Patterson & Hodges, 1992).

Our interest in this study is to establish more directly that different learning experiences with words affect subsequent encounters with those words. Specifically, using both behavioral measures and recordings of event-related potentials (ERPs) we want to examine the differences in word identification and meaning retrieval as a consequence of different word learning contexts. The benefit of using ERPs to examine the acquisition of word knowledge is that they provide information on the order of milliseconds on the time course of word processing, and they can potentially distinguish among the cognitive events that occur at word reading (e.g., recognizing a word as familiar versus recognizing a word as recently experienced).

In addition to using ERPs to expose the consequences of word experience, we also address the research question of individual differences in comprehension skill and word learning. Skilled reading depends on knowing many words in the language (Anderson & Freebody, 1981; Sternberg & Powell, 1983), and that the underlying meaning of these words can be easily and reliably accessed from the words' form (Perfetti, 1985). Word knowledge is a critical component of skilled reading as indicated by the Lexical Quality Hypothesis (Perfetti, 2007; Perfetti & Hart, 2001, 2002) and other studies (Daneman, 1984; Everatt & Underwood, 1994), and it may also support new word learning. If so, we would expect to observe differences in word learning as a result of individual differences in comprehension skill. These differences may be observed in ERP components that indicate recognition and meaning processing for recently experienced words.

1.1. What is 'known' about a word?

In this study, we define *word knowledge* as information that a reader has about a word's form (orthography and phonology)

* Corresponding author. Address: Learning Research and Development Center, Rm 641, 3939 O'hara Street, University of Pittsburgh, Pittsburgh, PA 15260, United States.
E-mail address: michal.balass@gmail.com (M. Balass).

0361-476X/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.cedpsych.2010.04.001

and its meaning. Meaning representations can be characterized by two distinct levels of knowledge. One is a specific and 'definition-like' representation of a word's meaning that is delimited by a set of core meaning features that are characteristic of a particular word (Drum & Konopak, 1987). This level of word representation is context-independent (e.g., Barsalou, 1982), and is referred to as *definitional* knowledge of a word (Chaffin, 1997). The second type of word information reflects general meta-linguistic knowledge of the word's relationship to other words (e.g., co-occurrence) and extendibility of use in other contexts (Miller, 1999). This level of word knowledge is context-dependent (e.g., Barsalou, 1982), and is referred to as *contextual* knowledge of a word (Chaffin, 1997; Miller, 1999; Miller & Charles, 1991). In this study, learners were exposed to the definitional knowledge of a word using dictionary definitions, and were later tested on the acquisition of this information using a word processing task.

1.2. How is a word learned?

Word knowledge can be communicated in many different ways (e.g., in context, dictionary definition, explicit vocabulary instruction, etc.). We argue here that regardless whether the word is learned 'incidentally' via reading (e.g., Herman, Anderson, Pearson, & Nagy, 1987; Jenkins, Stein, & Wysocki, 1984; Nagy, Herman, & Anderson, 1985), explicitly in a vocabulary instruction setting (Beck, McKeown, Omanson, 1987), or actively using derivational learning tasks (e.g., Fukkink & de Glopper, 1998; van Daalen-Kapteijns & Elshout-Mohr, 1981) all features of the encounter with the word are encoded in memory. In this way, the information that is retrieved about a word when it is encountered is the accumulation of all encoded learning experiences.

This view is the basis for the instance-based word learning framework (Reichle & Perfetti, 2003). According to this framework, learning is driven by a resonance process where word knowledge is acquired from information that is repeated (most frequent) across instances of word use. Reichle and Perfetti (2003) demonstrated that an instance-based resonance mechanism can account for both word frequency and regularity effects. For word learning, a resonance mechanism makes useful predictions on the effects of word experience. For example, Bolger, Balass, Landen, and Perfetti (2008) demonstrated that an instance-based model can account for the effects of sentence contexts and dictionary definitions on the acquisition of meaning.

According to the instance-based framework, each encounter with a word is encoded in specific episodic traces in memory. An episodic trace is a record of an encounter with a word that potentially preserves all the properties of that encounter. When a learner attends to the word, information about the word's basic features (e.g., orthography, phonology, and meaning) as well as contextual features of the word encounter is encoded in memory. Each encounter with a word gives rise to its own set of episodic traces. As a result, aspects of the encounter that are attended to and are recurring will be represented repeatedly in memory. In this way, repeated word features that are represented in multiple episodic traces will be strengthened with each encounter in which they have appeared. Eventually, this process creates a set of unified traces that become characteristic of the word's spelling, pronunciation, and meaning representation. Encoded episodic traces for any given word are retrieved by means of a resonance process that activates stored traces from memory when a word is encountered in a new context.

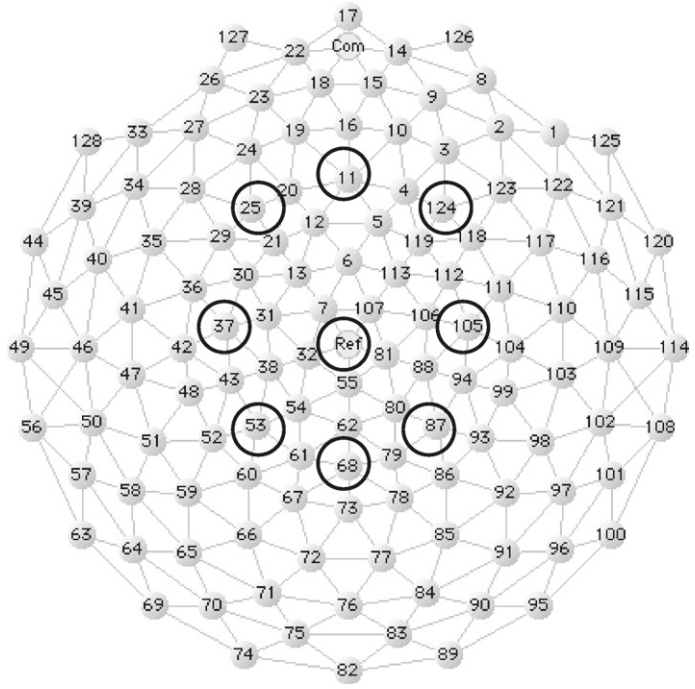
The instance-based approach makes useful predictions about the effects of word learning on two related processes; word recognition (Reichle and Perfetti (2003) refer to this as familiarity) and meaning acquisition. Given that each instance with a word is encoded in memory, different word experiences will give rise to episodic traces differing in strength. For example, in some word encounters only the word's spelling and pronunciation may be encoded, leading to weaker meaning traces. The accumulation of these traces across all word experiences for a particular word will affect how it is processed in a subsequent encounter. Under the assumption that word recognition is affected by memory strength (Finnigan, Humphreys, Dennis, & Geffen, 2002), words that are represented with stronger traces will be more familiar than words are represented by weaker memory traces. In the same way, the information that is acquired and retrieved of a word's meaning will depend on whether meaning features have been repeated or strengthened in multiple episodes with the word. Different word experiences will give rise to episodic traces differing in strength of the word's spelling, pronunciation, and meaning. The instance-based model provides some general hypotheses regarding the effects of experience that may affect word learning that can be tested by measuring ERPs.

1.3. ERPs and word learning

According to the instance-based framework, different word experiences will affect word learning and processing of subsequent encounters with the word. Behavioral measures alone by way of reaction times and accuracies provide information about learning once the processing of a particular word has been completed. Specifically, behavioral measures reflect information about processing of a learned word following the completion of sensory, cognitive, and motor processes (Bentin, 1989). For this reason, behavioral measures alone may not be adequate in revealing differences between different word learning experiences, especially in the early stages of word identification. For this current study, we are also interested in the cognitive events that happen initially at word reading as a consequence of differing word learning experiences. Thus, along with behavioral measures that can reveal whether word meaning has been learned, ERPs can be used to examine the consequences of these experiences to expose the cognitive events associated with word encounters (i.e., word recognition and meaning retrieval).

ERPs reflect recorded electroencephalogram (EEG) measurements from the scalp that are elicited in response to a particular stimulus (e.g., a word). One important advantage of ERPs is that they allow the researcher to measure cognitive activity online on the order of milliseconds for the entire process of word identification (i.e., recognition and meaning retrieval). An ERP record consists of multiple discrete components or brain waves that are related to different stages of stimulus processing. For example, word recognition processes are associated with specific ERP components and may not overlap with components associated with meaning retrieval processes. ERP components are distinguished in an ongoing EEG record by the direction of their amplitude, when they occur on the EEG record, and location on the scalp. Most ERP components are usually designated by the polarity of their amplitude in microvolts (positive, P or negative, N) and by the latency (the time when the component appears on the EEG record after the presentation of the stimulus) of the component's maximal amplitude in milliseconds (ms). For example, the N400 is an ERP component that is associated with semantic processing. It has a negative amplitude and it appears on the EEG record approximately 400 ms after the presentation of a word. Components are typically observed across different scalp locations. For instance, components associated with auditory processing may be observed in temporal sites, visual processing in occipital sites, and semantic processing in central and parietal sites on the scalp. Fig. 1 illustrates these different electrode locations on the scalp.

Several studies have indicated the benefits of using ERPs to examine word learning (Key, Molfese, & Ratajczak, 2006;



10-20 System	Electrode Number
F3	25
Fz	11
F4	124
C3	37
Cz (Ref)	129
C4	105
P3	53
Pz	68
P4	87

Fig. 1. Map of electrode sits on the scalp based on the 10–20 system. The table on the right shows the conversion between the 10–20 system notation and the electrode numbers. The isolated common is denoted by 'COM', and the reference is denoted by 'REF'.

McCandliss, Posner, & Givón, 1997; Perfetti, Wlotko, & Hart, 2005), familiarity (Curran, 1999; Finnigan et al., 2002; Widling & Rugg, 1997), meaning processing (Kutas & Hillyard, 1980), and word recognition (Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier, 1999). This previous research has identified several ERP components correlated with learning that are relevant to our hypotheses. One ERP component, also referred to as the 'old/new' effect or the P600, has been implicated in recognition processes. The other ERP component of interest, the N400, reflects meaning retrieval processes.

1.4. The old/new ERP effect: P600

The 'old/new' effect refers to an ERP recognition memory component (P600) that distinguishes between recently presented items and unpresented items. The P600 effect is a positive going waveform with central and parietal electrode distribution on the scalp that appears around 500–800 ms after the onset of a word. The P600 is characterized by more positive amplitudes for 'old' presented items than for 'new' unpresented items (Curran, 1999; Rugg, Allan, & Birch, 2000). The P600 old/new effect has also been directly observed in word learning studies. For example, Perfetti et al. (2005) exposed learners to the form and meaning of rare unknown words. They reported a positive component that peaked at around 500 ms (i.e., P600) after the presentation of a word, and showed larger amplitudes after the presentation of a learned word than unpresented rare or familiar words. Thus, they took the P600 to be a marker for a recently learned word.

1.5. Meaning acquisition and processing effects: N400

The N400 component is related to meaning processing. It is a negative going waveform that appears from 300 to 500 ms after the onset of a word, and its distribution is also at central and parietal sites. The N400 has been implicated in many tasks requiring semantic processing. Generally, the N400 reflects meaning congruence between a word and its previous context. It has been implicated in both sentence context presentations (Kutas & Hillyard, 1980) and single word semantic priming presentations (Nobre & McCarthy, 1994). When a word is incongruent with its context it produces a negative-going wave peaking at about 400 ms after the onset of a word; a congruent word produces a reduced N400 effect (less negative amplitude). Perfetti et al. (2005) reported N400 effects related to meaning acquisition. Following word training in their study, learners were presented with the trained words followed by meaning probes that were related or unrelated to the words' meaning. They observed robust N400 effects for unrelated meaning probes and a reduction of the effect for related meaning probes. The reduction of the N400 effect for related probes indicated that participants learned the meanings of the words.

1.6. Individual differences

General differences in reading ability and comprehension have been shown to be contributing factors to word learning (Cain, Lemmon, & Oakhill, 2004; Curtis, 1987; Jenkins et al., 1984; McKeown, 1985). McKeown (1985) and van Daalen-Kapteijns and Elshout-Mohr (1981) found that skilled readers were better able to make use of a word experience to generate a hypothesis about its meaning than less-skilled readers. Daneman (1984) proposed that word knowledge and comprehension skill reflect a general ability to acquire new information. Skilled readers are successful at learning words from different types of word experiences because they can incorporate already known word knowledge to assist themselves in learning learn new words.

Perfetti et al. (2005) reported that differences in comprehension skill and word learning are detectable in ERPs. In their study, skilled readers learned slightly more words than less-skilled readers, and showed stronger P600 effects (i.e., greater amplitudes) for learned words. Also, skilled readers showed superior performance on meaning judgments compared with less-skilled readers. Perfetti et al. (2005) attributed these differences in learning to differences in the process of encoding episodic traces. Skilled readers may make better use of their experiences with words, laying down

stronger episodic traces for form and meaning, resulting in better word learning.

In the present study, the episodic word experience model leads to the hypothesis that differences in word experience will give rise to encoded memory traces that differ in strength of a word's spelling, pronunciation, and meaning; this will affect subsequent word identification and meaning retrieval processes. To test this hypothesis we exposed learners to a set of unknown rare English words in three learning conditions. The available information for these words varied such that a learner was exposed to two of the three constituents of a word (i.e., orthography-meaning, orthography-pronunciation, phonology-meaning), thus varying the type of experience with the words. We predicted that following learning, during a semantic-judgment task, we would observe differences in processing as a consequence of word experience: ERP evidence that learners became familiar with and acquired the meaning of the words; a P600 effect that would show differences between learned words and unlearned words; and an increased N400 for words for which they did not know the meaning. Based on previous evidence, we expected that high-skilled readers would show more effective learning than less-skilled readers. This difference would be observed by ERP differences in recognition processes and meaning retrieval as indicated by the P600 and N400 components.

2. Methods

2.1. Participants

Learners were recruited from a pool of approximately 500 college-age students from the University of Pittsburgh department of Psychology undergraduate participant pool. All participants in this pool completed a battery of tests assessing comprehension ability and lexical knowledge for course credit. From this larger pool, 37 right-handed (females = 21, males = 16) individuals were recruited (via email) to represent a range of reading comprehension skill as assessed by the Nelson–Denny comprehension test (Nelson & Denny, 1973). Their participation was completely voluntary and did not satisfy any requirements related to coursework. Participants' handedness was determined using a short questionnaire to assess their handedness preferences. Three learners were eliminated from the final analysis due to unusable EEG data. Learners were paid $24 for completing the 3-h study.

2.2. Comprehension skill assessment

Participants for this study were recruited for participation based on their Nelson–Denny comprehension test score (Nelson & Denny, 1973). Previous studies (Hart, 2005; Landi, 2005), and more recently, Nelson and Halderman (2008), have shown significant correlations between Nelson–Denny comprehension scores, lexical knowledge, and measures of non-verbal intelligence. Specifically, Landi (2005) showed significant correlations between comprehension, vocabulary, and spelling in an analysis of over 700 individuals who completed the Nelson–Denny test and other measures of vocabulary, spelling, and decoding skills. She reported a correlation of $r = 0.664$, $p < .01$ between comprehension and vocabulary, and a correlation of $r = 0.244$, $p < .01$ between comprehension and spelling. Nelson and Halderman (2008) reported similar correlations in an analysis of 2123 individuals who completed the same battery of tests in addition to the Raven Progressive Matrices (Raven, 1960). They reported a correlation of $r = 0.159$, $p < .01$ between comprehension and Raven's test. We have taken these results to indicate that the Nelson–Denny comprehension test is significantly correlated with other types of reading skills (i.e., vocabulary and spelling) and non-verbal skills (Raven's test), and therefore, it is a critical measure to use for recruiting participants. Thus, given that we were interested in testing comprehension skill differences in word learning, we used comprehension score, which may also account for other differences in reading skill, as the basis for recruiting participants.

We administered the Nelson–Denny test as a timed test. Participants had 15 min to read several passages and to attempt to answer 36 multiple choice comprehension questions. Each multiple-choice question had five choices. The test was scored in such a way that participants received two scores on their performance; an accuracy score (how many questions were answered correctly) and a speed score (how many questions out of 36 questions were attempted). These two scores were combined into one composite comprehension score; the number of correct items was subtracted from the number of attempted items, and then this number was then multiplied by 0.20. We employed a fifth of a point penalty for every question that was answered incorrectly. Thus, our composite comprehension score formula was as such: comprehension score = $0.20 * (\text{attempted items} - \text{correct items})$. This scoring method is not the standard protocol for scoring the test (e.g., as suggested by Cummins, 1981); in our measure we combined accuracy and speed for one composite score, whereas standard scoring procedures suggest separate scores for accuracy and reading speed.

Individuals whose composite score was 22.2 or lower were considered to be less-skilled readers, and those whose composite score was 22.2 or higher were considered to be high-skilled readers. This criterion was determined based on a 50–50 median split of the distribution of scores from our sample of 37 participants. This median score is comparable to other college-age samples. One such sample reported a mean of 20.36 ($SD = 5.40$) for a 15-min Nelson–Denny administration with a sample of 307 participants (Creaser, Jacobs, Zaccaria, & Carsello, 1970). More recently, Nelson and Halderman (2008) reported a mean of 18.18 ($SD = 6.99$) for the same Nelson–Denny 15-min procedure with a larger sample of 1453 participants. Carver (1992) used a 20-min Nelson–Denny procedure, and reported a similar mean of 21.9 ($SD = 5.0$) with a sample of 65 participants. This evidence suggests that our smaller study sample is representative of the typical scores attained on the Nelson–Denny comprehension test. Specifically, in our sample, the overall comprehension score was $M = 22.36$, $SD = 5.19$, with a minimum score of 9.6 to a maximum score of 30.4. High-skilled readers' scores ranged from a minimum score of 22.4 to a maximum score of 30.4, and $M = 26.60$, $SD = 2.46$. Less-skilled readers' scores ranged from a minimum score of 9.6 to a maximum score of 22.2, and $M = 18.18$ and $SD = 3.51$.

2.3. Procedure

The experiment included two phases; word learning and testing. During the approximately 2-h long learning phase, rare words were presented to the participant under three learning conditions on a computer screen using E-prime software from Psychological Software Tools, Inc. If the participant failed to learn all of the words to criterion of 100% after 2.5 h, then they were compensated for their time without completing the testing phase of the experiment. Only participants who were successful in reaching the learning criterion completed the testing phase of the experiment. No participant was eliminated due to failure to complete the learning phase in the allotted time. The testing phase of the experiment which included the recording of ERPs and the completion of a semantic-judgment task was approximately an hour long. The entire experiment duration ranged between 3 and 3.5 h.

2.4. Materials

The learning phase of the experiment included a set of 105 rare English words, with 35 words randomly assigned to each of the three training conditions. The testing phase included an additional 35 unpresented rare words (we refer to these as "unlearned"), 35 unpresented familiar low-frequency words, and 35 unpresented filler words. Filler words were high-frequency words used to balance the range of word frequency of our stimuli used in the testing phase of the experiment. This was to ensure that participants could respond correctly to a portion of the words presented in the testing phase. These filler words were not of interest to our hypothesis, and therefore excluded from any analyses. Thus, the experiment included a total of 140 rare English words of which 105 were presented during the learning phase, 35 familiar low-frequency words, and 35 high-frequency filler words. Of the 140 rare words, only 11 appeared in the English Lexicon Project (Balota et al., 2007), and seven of those were in the Kucera and Francis database (Kucera & Francis, 1967). The mean word length of the rare words was 7.13 letters (range 4–9) and 2.49 syllables (range 1–4). Of the 35 familiar low-frequency words, 34 appeared in the English Lexicon Project, and 24 of those were in the Kucera–Francis database (mean frequency of 9.72; range 1–38). The mean word length of the familiar low-frequency words was 7.22 letters (range 4–9) and 2.57 syllables (range 2–3). Of the 35 filler words, 35 appeared in the English Lexicon Project, and 32 of those were in the Kucera and Francis database (mean frequency of 33.94; range 1–369). The mean word length of the filler words was 5.89 letters (range 4–12) and 1.66 syllables (range 1–4). The entire set of materials for this study can be found in Appendix A.

2.5. Word learning phase

Participants were presented with a total 105 rare words, with 35 words assigned to each of the three learning conditions described below. All learners experienced all three learning conditions. Each learning condition block began with an initial familiarity check (details are described in the following section). To control for order and sequence effects, a Latin square was used to counterbalance the learning conditions across participants; this procedure yielded six different learning conditions orders. A one-way ANOVA did not indicate any significant differences in the order or sequence of the learning conditions for reaction time performance, $F(5, 27) = 0.171$, $p = .971$ or accuracy performance, $F(5, 27) = 0.326$, $p = .893$. Words within each of the learning conditions were randomly assigned to eliminate any order effects. The learning phase was self-paced so the learner could spend as much time as needed on any given word. The learning conditions were as follows.

2.5.1. OM condition (orthography-to-meaning)

In this condition, participants were required to learn the spelling of the word along with its meaning. They were presented with the spelling of the word, its definition, and an example sentence using the word in context. No explicit information on the word's pronunciation was provided.

2.5.2. OP condition (orthography-to-phonology)

In this condition, participants were required to learn the pronunciation of the word along with its spelling. They were presented with the spelling of the word and heard its pronunciation. Participants were able to repeatedly hear the pronunciation of the word as many times as needed by pressing the appropriate key. No explicit information on the word's meaning was presented.

2.5.3. PM condition (phonology-to-meaning)

In this condition, participants were required to learn the meaning of the word in addition to its pronunciation. They heard the word along with a written presentation of its definition and an example sentence with the target word replaced by a blank space. Participants were able to repeat the pronunciation of the word as many times as needed by pressing the appropriate key. No explicit information on the word's spelling was provided.

Learners were required to learn all of the words to a 100% criterion. This criterion was imposed to ensure that the differences observed across comprehension skill were not due to lack of time to learn all of the words. The learning phase ended either when the criterion was reached or after 2.5 h had elapsed. As mentioned previously, if the participant was not able to learn all the words in the time allotted, they did not continue onto the testing phase. This time restriction was imposed based on pilot data that indicated that approximately 2 h were sufficient to learn these words to criterion regardless of comprehension skill.

2.6. Initial familiarity check

The learning phase of the experiment included a familiarity check to assess previous familiarity with the rare words. Regardless of the learning conditions order, learning in each condition began with a set of instructions followed by the familiarity check. For each block of 35 words within each condition, participants proceeded once through each word and responded whether it was familiar or unfamiliar by pressing the appropriate keys. A time limit of 4 s was imposed for the maximum decision time on each word. This time limit was used to ensure that participants made their responses quickly and made their decisions uniformly without fixating on any given word. If the participant responded within 4 s after the word appeared on the screen, they were presented with the relevant word information for that condition, (i.e., orthography-meaning, orthography-pronunciation, or phonology-meaning) and were able to study each word as long as needed. If 4 s elapsed without a response, the participant was not given the opportunity to study the word, and the next word appeared on the screen.

2.7. Reaching criterion

After completion of the familiarity check and initial learning phase, the participants were presented with the words again. Words were randomized but presented in the same block order established in the familiarity check. Participants saw or heard the target word and were asked to respond verbally with the appropriate definition or pronunciation of the word depending on the learning condition. A trained experimenter judged the accuracy of the response and noted whether it was correct or incorrect. The experimenter provided feedback by informing the participant of their accuracy. Regardless of the accuracy of the response, the participant was given the opportunity to study the word again for as long as they wanted before proceeding to the next word.

This procedure was repeated until the participant learned 100% of the words. When the participant provided the correct definition or pronunciation for any given word twice in a row, the word was considered learned and was removed from the list and did not appear again for the remainder of the learning phase. The learning criterion was reached when learners provided accurate responses for 100% of the words twice in a row. In this way, each word was encountered as least three times; once during the familiarity check and at least twice to reach the learning criterion.

If the learning criterion was reached in the allotted time of 2.5 h, the participant proceeded to the testing phase of the study where their EEGs were recorded. Otherwise, they were

compensated for their time spent during the learning phase of the experiment.

2.8. Testing phase

2.8.1. Semantic-relatedness judgment task

Upon successful completion of the learning phase, participants completed a semantic-relatedness judgment task to test their newly acquired word knowledge while their ERPs were recorded. Participants were presented with a single word followed by a second word (a meaning probe) that was related or unrelated in meaning. They were asked to respond whether the two words were related in meaning by pressing the appropriate keys. The presentation of the first word was either a word from the learning phase, an already known familiar low frequency word, a filler word, or a rare unlearned word. The trials were split into two halves with the meaning probe semantically related to the first word ("YES" trials; *gloaming-evening*) in one half and unrelated ("NO" trials; *wyvern-spear*) in the other half. The semantically related meaning probes were created by the experimenter, and were not derived directly from the definition the participant studied during the learning phase. That is, the experimenter avoided using any of the words used to define the word in the learning phase as related meaning probes. Semantically unrelated pairs were created by shuffling the meaning probe word pairs and randomly making new pairings. The order of the word pairs was randomized for each participant. Each learner completed a set of 210 trials while their ERPs were recorded.

Each trial began with a fixation cross that appeared in the middle of the computer screen for 350 ms. Following fixation, a blank screen was presented for a variable duration of 50–250 ms. The purpose of the variability in duration of the blank screen was to reduce any influence of non-stimulus related time-locked electrical activity. Following the blank screen, a prime word was selected randomly from the total set of words available for the task (a learned word from any of the three learning conditions, a familiar word, or an unfamiliar rare word). The first word was presented for 1000 ms and was followed immediately by the meaning probe presented for a maximum of 2000 ms.

Learners were instructed to press the '1' key with their right index finger if the two words presented were related in meaning and to press the '2' key with their right middle finger if the two words were unrelated in meaning. The meaning probe word was removed from the screen when a response was made or if no response was detected after 2000 ms had elapsed. Participants received feedback on the correctness of their judgment for each trial. If the participant was correct, the word 'correct' in green font color and the decision time were displayed on the screen for 2000 ms before the onset of the next trial. If the response was incorrect, the word 'incorrect' in red font color was displayed with no decision time.

2.9. ERP recordings

Before beginning the semantic-judgment task, participants were fitted with an electrode cap. The scalp potentials were recorded from 128 sites using a Geodesic Senor Net (Electrical Geodesics, Eugene, OR) with Ag/AgCl electrodes. All impedances were kept under 40 kΩ. A vertex reference was used in the recording, and the data were recomputed offline against the average reference. Six eye channels allowed rejection of trials with eye movements and eye blink artifacts. The EEG signals were recorded with a sampling rate of 1000 Hz. The hardware filter was between 0.1 and 200 Hz. A 30 Hz lowpass filter was applied to all learners' data. ERPs were stimulus-locked to two separate events; the first word and the second word (i.e., the meaning probe). For each event, ERPs were averaged over a 1100 ms time segment; 100 ms baseline and a 1000 ms epoch that was either defined by the presentation of the first word or the second word (i.e., the meaning probe). Bad channels were removed from the recordings and replaced by spherical spline interpolation using data from the remaining channels. This method has been shown to be highly effective in approximating the average surface potential on the scalp (e.g., Ferree, 2006; Scherg, Ille, Bornfleth, & Berg, 2002). Any trials containing eye-movements, eye blinks, or channel artifacts were removed and not used in the final analyses. For each participant, no more than five trial segments were rejected due to eye-movements or eye blinks. If more than 5 trial segments were rejected due to eye-movement artifacts, the participant's data were excluded from the analyses. On average, for each participant, we retained 30 trials per condition after artifact detection. Following trial rejection, ERPs were transformed using average reference and then corrected using a 100 ms baseline.

3. Results

3.1. Behavioral results

The decision time and accuracy results for the semantic-judgment task are shown in Tables 1 and 2. As indicated by Table 1, decision times varied between 650 and 950 ms depending on the word type (OM, OP, PM, familiar, unlearned), relatedness (related, unrelated), and correctness (correct, incorrect) of response. Decision times for learned (OM, OP, PM) and familiar words were faster than for unlearned rare words, decision times for related words were slightly faster than unrelated words, and decision times for correct responses were faster than incorrect responses. An analysis of variance of word type (OM, OP, PM, familiar, unlearned) × relatedness (related, unrelated) × correctness (correct, incorrect) with a between-subjects factor of comprehension skill (high-skilled, less-skilled) for the decision times indicated a significant main effect for word type $F(4, 29) = 23.95$, $p < .01$, correctness $F(1, 32) = 19.70$, $p < .01$, and no significant main effect for relatedness. However, a significant word type × relatedness interaction $F(4, 29) = 3.21$, $p < .05$ indicated that learners were faster for related trials in the OM and PM learning conditions. An interaction of word type × correctness, $F(4, 29) = 7.62$, $p < .01$ showed that decision time differences were present for OM, PM, and familiar words only, with OP words and unfamiliar words showing no differences. No significant main effects or interactions were observed for comprehension skill.

As shown in Table 2, learners were more accurate in their responses for OM, PM, and familiar words than OP and unlearned words, and showed higher overall accuracy for 'no' unrelated trials. An analysis of variance of word type (OM, OP, PM, familiar, unlearned) × relatedness (related, unrelated) with a between-subjects factor of skill (high-skilled, less-skilled) indicated a significant main effects for word type $F(4, 29) = 121.0$, $p < .01$, relatedness $F(1, 32) = 37.08$, $p < .01$, and a word type × relatedness interaction $F(4, 29) = 23.72$, $p < .01$. The interaction showed a 'no' response bias for OP and rare unlearned words. This is a tendency for learners to judge word pairs for which they do not know the meaning (i.e., OP words or unlearned rare words) as unrelated to its meaning probe. No significant main effect or interactions were observed for comprehension skill.

3.2. Event-related potentials results

ERPs were analyzed using the two different time windows appropriate for the N400 and P600, 300–500 ms and 500–700 ms, respectively. Data were averaged at each of the nine standard electrode sites based on the 10–20 electrode placement system (F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4). Each electrode site

Table 1
Behavioral results: semantic-relatedness decisions times (in milliseconds).

Condition	Skilled			Less skilled			
	Correct	Incorrect	*M*	Correct	Incorrect	*M*	*M*
OM							
Related (YES)	649.85 (71.2)	774.71 (133)	*712.28*	716.67 (111)	723.71 (136)	*720.19*	*716.24*
Unrelated (NO)	711.01 (83.0)	776.68 (180)	*743.85*	758.07 (149)	804.52 (178)	*781.30*	*762.57*
M	*680.43*	*775.70*	*728.06*	*737.37*	*763.76*	*750.74*	*739.40*
OP							
Related (YES)	862.33 (191)	832.79 (114)	*847.56*	861.16 (163)	852.44 (184)	*856.80*	*852.18*
Unrelated (NO)	843.82 (108)	847.63 (148)	*845.73*	864.37 (183)	936.47 (191)	*900.42*	*873.07*
M	*853.08*	*840.21*	*846.64*	*862.77*	*894.46*	*878.61*	*862.63*
PM							
Related (YES)	725.72 (106)	820.36 (155)	*773.04*	759.88 (147)	876.60 (204)	*818.24*	*795.64*
Unrelated (NO)	757.43 (124)	847.27 (158)	*802.35*	814.80 (159)	853.43 (179)	*834.12*	*818.23*
M	*741.58*	*833.82*	*787.70*	*787.34*	*865.02*	*826.18*	*806.94*
Familiar							
Related (YES)	732.09 (72.5)	940.98 (204)	*836.54*	740.21 (114)	896.38 (220)	*818.30*	*827.42*
Unrelated (NO)	772.75 (90.8)	827.63 (153)	*800.19*	782.48 (132)	855.22 (189)	*818.85*	*809.52*
M	*752.43*	*884.31*	*818.36*	*761.35*	*875.80*	*818.57*	*818.47*
Unlearned rare							
Related (YES)	927.62 (140)	825.44 (115)	*876.53*	971.74 (174)	838.33 (173)	*905.04*	*890.78*
Unrelated (NO)	829.26 (103)	929.45 (162)	*879.36*	835.91 (160)	875.01 (172)	*855.46*	*867.41*
M	*878.44*	*877.45*	*877.94*	*903.83*	*856.67*	*880.25*	*879.10*
Overall *M*	*781.19*	*842.29*	*811.74*	*810.53*	*851.23*	*830.87*	*821.31*

Table 2
Behavioral results: percent accuracy.

Condition	Skilled	Less skilled	*M*
OM			
Related (YES)	88.3 (.077)	89.2 (.078)	*88.6*
Unrelated (NO)	87.0 (.088)	81.5 (.097)	*84.3*
M	*87.7*	*85.4*	*86.5*
OP			
Related (YES)	41.0 (.208)	39.5 (.177)	*40.1*
Unrelated (NO)	75.3 (.148)	73.1 (.208)	*74.2*
M	*58.2*	*56.3*	*57.2*
PM			
Related (YES)	79.3 (.094)	75.8 (.176)	*77.6*
Unrelated (NO)	82.4 (.123)	79.1 (.138)	*80.6*
M	*80.9*	*77.5*	*79.1*
Familiar			
Related (YES)	93.1 (.063)	82.9 (.170)	*88.0*
Unrelated (NO)	89.1 (.057)	86.3 (.100)	*87.7*
M	*91.1*	*84.6*	*87.9*
Unlearned rare			
Related (YES)	28.8 (.216)	30.3 (.208)	*29.6*
Unrelated (NO)	78.9 (.165)	80.0 (.156)	*79.5*
M	*53.9*	*55.2*	*54.5*
Overall *M*	*74.32*	*71.77*	*73.02*

has a letter and a number to identify its location on the scalp. The letters in this notation (e.g., F, C, and P) signify the lobe and the numbers signify the hemisphere (i.e., odd numbers are left hemisphere locations and even numbers are right hemisphere locations; z denotes midline electrodes). Thus, electrodes F3, Fz, and F4 are located on frontal locations on the scalp; electrodes C3, Cz, and C4 are located on central locations on the scalp; electrodes P3, Pz, and P4 are located further back, on parietal locations on the scalp. Fig. 1 is a graphical depiction of these electrode sites on the scalp. These nine electrodes were selected based on previous research indicating that P600 effects are typically observed in central and parietal electrode sites (Curran, 1999; Rugg et al., 2000) and N400 effects are typically observed in central electrode sites (Kutas & Hillyard, 1980). Further, frontal electrodes were used in these analyses to maintain proper convention, and as comparison electrodes in which we did not expect to observe a P600 or N400 effect. Based on previous findings, we hypothesized that we would observe a P600 at about 600 ms after the presentation of the first word and that the amplitude of this component would be modulated by word experience (e.g., presented words from the learning phase would show greater positive amplitudes than unpresented words). Also, we expected to observe a meaning effect indicated by the N400 after the presentation of the second word (i.e., meaning probe); the amplitude of this component would be reduced for related meaning probes. Thus, the time windows used in the analyses were selected to span 100 ms in each direction of the hypothesized time in which we expected to observe the ERP effects.

Mean amplitude values were computed across all time points in the specified time window. Two analyses were completed: 500–700 ms after the presentation of the first word and 300–500 ms after the presentation of the second word (i.e., the meaning probe). A word type (OM, OP, PM, familiar, unlearned) × lobe (frontal, central, parietal) × hemisphere (left, midline, right) ANOVA with a between-subjects factor of skill (high-skilled, less-skilled) was completed to test for the P600 effect. To test for significant N400 effects, we added an additional factor of relatedness (related, unrelated) yielding a word type (OM, OP, PM, familiar, unlearned) × relatedness (related, unrelated) × lobe (frontal, central, parietal) × hemisphere (left, midline, right) ANOVA with a between-subjects factor of skill (skilled, less-skilled). Greenhouse–Geisser (*GG*) correction (Greenhouse & Geisser, 1959) was applied to all factor levels that violated sphericity. Mean amplitudes for the P600 and the N400 are reported in microvolts (mV).

3.3. P600: 500–700 ms after the presentation of the first word

An analysis of variance of the mean amplitude showed a significant word type × lobe interaction $F(8, 25) = 2.57$, $p < .05$, *GG* $\varepsilon = 0.452$ and a word type × lobe × skill interaction $F(8, 25) = 3.78$, $p < .01$, *GG* $\varepsilon = 0.452$. More positive amplitudes were observed in parietal electrodes for learned words (OM = −.036, OP = −.381, PM = −.594) than for familiar (−.828) and unlearned (−.702) words. As shown in Fig. 2, high-skilled readers showed more positive amplitudes for presented learned words (OM, OP, and PM) than for unpresented familiar and unlearned words. Less-skilled

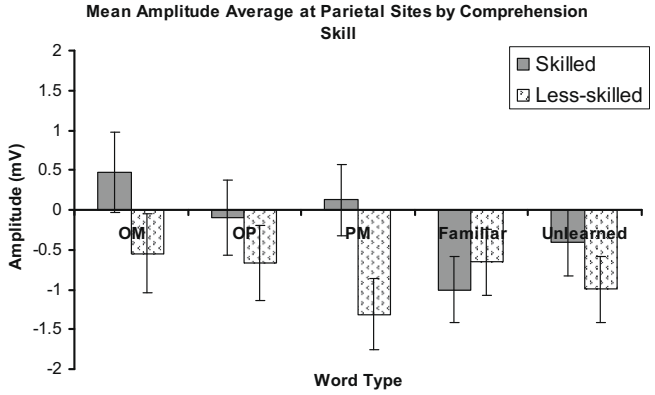


Fig. 2. Mean amplitude effects for the P600 illustrating the amplitude differences by word type and comprehension skill.

readers did not show any differences between any of the word types. Also, a comprehension skill difference was observed for the PM condition; high-skilled readers showed more positive amplitudes than less-skilled readers, $t(32) = 2.25$, $p < .05$.

3.4. N400: 300–500 ms after presentation of the meaning probe

A word type (OM, OP, PM, familiar, and unlearned) × relatedness (related, unrelated) × lobe (frontal, central, parietal) × hemisphere (left, midline, right) ANOVA was conducted on the mean amplitude after the presentation of the second word (i.e., meaning probe). Results showed a main effect of word type $F(4, 29) = 9.86$, $p < .01$. More positive amplitudes were observed for OM (.431), PM (.059), and familiar words (.134) than for OP (−.237) and unlearned (−.368) words. A main effect of relatedness was observed as well, $F(1, 32) = 17.60$, $p < .01$. More positive amplitudes were observed for related, "YES" trials (.120) than unrelated, "NO" trials (−.113). A word type × relatedness interaction, $F(4, 29) = 5.78$, $p < .01$, revealed more positive amplitudes for related ("YES") trials for the OM, OP, and unpresented familiar words than for unrelated trials ("NO"). This effect of relatedness was not observed for OP and unlearned words as illustrated in the bar graph in Fig. 3.

Based on the results of the word type × relatedness interaction, we chose to illustrate the effects of word type and relatedness at the Cz electrode in Fig. 4. The Cz electrode is a good candidate for showing these effects because of previous evidence (e.g., Kutas & Hillyard, 1980) showing clear N400 effects at central electrode sites. As mentioned previously, greater negative amplitudes were observed for OP and unlearned rare words than for OM, PM, and familiar words. Furthermore, a reduction of the N400 (i.e., greater positive amplitudes) was observed for related trials. This effect was

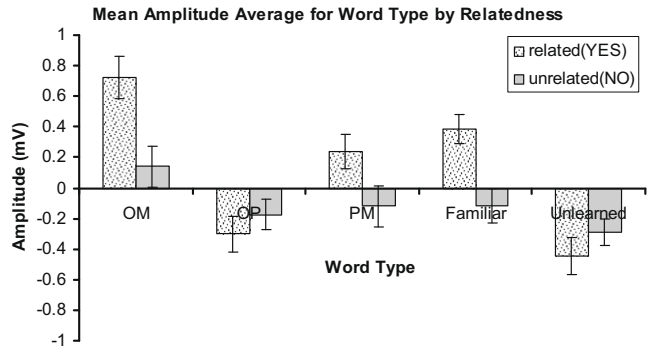


Fig. 3. Mean amplitude effects for word type by relatedness ("YES" trials versus "NO" trials).

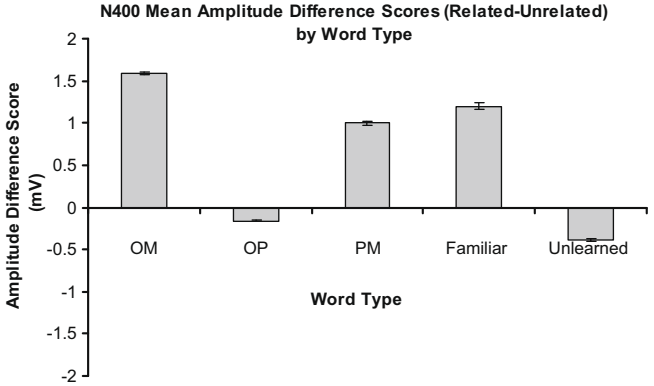


Fig. 4. Difference amplitude scores for word type by relatedness.

modulated by the word type, however. As illustrated in Fig. 4, the magnitude of the amplitude difference between related, ("YES") trials versus unrelated ("NO") trials depended on the word type. To illustrate this effect, the difference in amplitude was calculated first by averaging the mean amplitude for word type by relatedness from 300 to 500 ms at the Cz electrode site. Then, the difference score was computed by subtracting the mean for unrelated trials from related trials. The largest magnitude difference was observed for OM (1.59) words, followed by familiar (1.20), and PM (0.999) words. Differences in amplitude for relatedness were not found for OP (−.159) and unlearned (−.381) words. To illustrate the difference in N400 amplitude for related versus unrelated trials, Fig. 5 shows a comparison of the OM words and unlearned words; the latency of interest is between 300 and 500 ms (ms), and it has been highlighted with a rectangle. The four waveforms represent OM related (YES) trials, OM unrelated (NO) trials, unlearned (YES) trials, and unlearned (NO) trials. The figure demonstrates two important effects; first, that there is a large difference between the amplitude of OM (YES) trials and OM (NO) trials, and second, that the difference between the amplitude of unlearned (YES) trials and unlearned (NO) trials is small. The more positive amplitude for OM (YES) trials showed a reduction in the N400 effect to which we interpret to indicate that our participants learned the meanings of the OM words. This effect can be compared with the unlearned (YES) trials to which we do not observe a reduction of the N400 because our participants did not have access to the meaning information for these words.

Additionally, we observed a skill difference for this relatedness effect. A relatedness × hemisphere × skill interaction, $F(2, 31)$, $p < .05$ showed comprehension skill differences in amplitudes for related, "YES" trials versus unrelated, "NO" trials. High-skilled readers showed a significant difference in amplitude for related (0.362) versus unrelated (−0.161) trials at midline locations, $t(16) = 3.71$, $p < .01$. Less-skilled readers did not show an amplitude difference for related (.333) versus unrelated trials (.555), $t(16) = .346$, $p = .734$.

4. Discussion

Our aim for this study was to investigate differences in word learning as a result of type of word experience, and to address any possible differences word learning across different groups of readers. The consequences of different types of word experience were observed both in our behavioral and ERP results. Further, differences in comprehension skill and word experience were observed in our ERP results. Learning effects were observed when learners made semantic judgments on related (e.g., *gloaming-evening*) and unrelated (e.g., *wyvern-spear*) word pairs, and in the ERP record after the first word and second word (i.e., meaning probe)

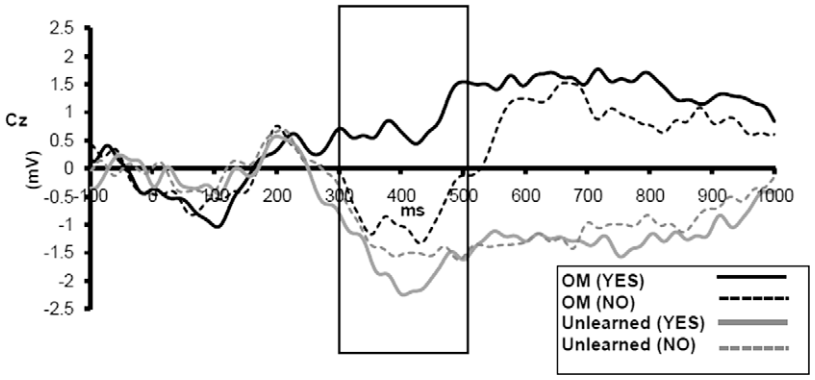


Fig. 5. Grand average waveform for the Cz electrode illustrating the N400 reduction effect for OM related words.

presentations. The behavioral results indicated that learners made faster and more accurate meaning decisions for presented learned words (OM, OP, and PM) than for unlearned rare words. No reaction time and accuracy differences were shown for OM, PM, and familiar words. Within the learning conditions, we observed the effects of word experience; learners were faster and more accurate for words for which they learned the meaning (i.e., OM and PM). No differences of comprehension skill were detected in the behavioral results.

The lack of evidence showing a comprehension skill difference in the behavioral data may be due to two factors; the learning criterion and sensitivity of reaction time and accuracy measures to differences in comprehension skill. Regardless of comprehension skill level, all participants learned the words to a 100% criterion before they proceeded to the testing phase of the experiment. As a result, all participants demonstrated equal knowledge of the words in the allotted time they had to learn them. This learning criterion may have benefitted the less-skilled and helped boost their performance on the semantic-judgment task. It may also be the case that the reaction time and accuracy measures alone were not sensitive enough to show differences in word learning between the two groups of readers. That is, behavioral measures may not be adequate in illustrating subtle differences in the quality of encoded word representations across different groups of readers. Using ERP methodology has a distinct advantage for such a problem given that this method allows us to examine the very early stages of word processing where comprehension differences may occur.

The ERP results indicated differences in two important processes related to word identification as a consequence of word experience: recognition of a previously experienced word and meaning retrieval. Word experience effects of the presented words during learning were observed approximately 600 ms after the first word was presented. Previous word learning literature has referred to this as an episodic training effect (Perfetti et al., 2005) that is similar to the old/new P600 effect observed in recognition memory paradigms (Curran, 1999; Rugg et al., 2000). Overall, learners showed greater amplitudes for experienced words than familiar known words and unlearned words which were not presented during the learning phase of the experiment. This shows that learners distinguished between presented and unpresented words (an episodic effect). In particular, this episodic effect was stronger for OM words than for OP and PM words as indicated by greater positive amplitudes for OM words (a word experience effect). As shown in Fig. 2, only skilled readers showed these effects; less-skilled readers' data suggest that they did not distinguish among presented and unpresented words (i.e., no episodic effect) or learning conditions (i.e., no word experience effect).

The modulation of the P600 amplitude as a result of word experience implies that orthographic and meaning components are especially important for word recognition. As indicated by the P600 amplitudes for OM, OP, and PM words, orthography-to-phonology links do not establish strong episodic traces in memory. Furthermore, visually presented words do not elicit recognition from phonology-only presentations even when meaning information is available (i.e., PM words). As illustrated in Fig. 2, high-skilled readers were able to benefit more from phonology-meaning information than less-skilled readers, but orthography-meaning information seems to be more critical for recognition. Thus according to our episodic word experience model (Reichle & Perfetti, 2003), these findings suggest that knowledge of a word's orthography-to-meaning (OM condition) is encoded in stronger memory traces than orthography-to-phonology (OP condition) and phonology-to-meaning (PM condition); this is evidenced by stronger episodic effects (i.e., bigger P600) for orthography-to-meaning (OM condition) words.

Our data also indicated that meaning retrieval was affected by word experience. The N400 effect was modulated by relatedness; greater negative amplitudes were observed for unrelated, "NO" trials than for related, "YES" trials. These findings replicate previous results (Kutas & Hillyard, 1980; Nobre & McCarthy, 1994; Perfetti et al., 2005) that report a reduction in N400 for words that fit their previous context (e.g., a related meaning probe). Like the P600 effect, the N400 varied by the type of word experience. A larger reduction of the N400 for related trials was observed for OM and PM words as illustrated in the difference scores plot in Fig. 4. The magnitude of this reduction for OP and unlearned words was slight. In particular, the amplitude differences between related and unrelated OM words were larger than PM words, and comparable to familiar words.

The differences in N400 amplitude for related and unrelated trials suggest that orthographic information is not only important for recognition, but is also critical for meaning retrieval. Specifically, as indicated by the difference amplitude scores for OM and PM words, meaning information was more reliably accessed from representational links to orthographic information than phonological information. These findings replicate previous results by Nelson, Balass, and Perfetti (2005) who found that orthographic learning leads to a better representation of word form than phonological learning. It may also be the case, as illustrated by our results that orthography establishes a stronger representational link to meaning, and that phonology alone may not be sufficient to establish a reliable episodic association to meaning.

Thus, the P600 and N400 results, along with our behavioral measures indicate that word experience affects recognition and meaning retrieval processes. As shown in our results, there is a distinct benefit for learning new word meaning along with orthographic information than phonological information. According to many hypotheses of vocabulary learning (e.g., Herman et al., 1987; Jenkins et al., 1984; Nagy et al., 1985), word knowledge is gained "incrementally" whereby information of a word's meaning

along with information of its form is learned through multiple and often variable encounters. As pointed out by McKeown (1985), the incremental accumulation of word knowledge (especially meaning) requires learners to recognize that they have experienced a word previously in a particular context and retrieve any associated meaning information. Over time, this process allows these word encounters (encoded episodic traces) to become a unified representation of the word. Evidence from our study suggests that the availability of repeated orthographic knowledge rather than repeated phonological knowledge may be more facilitatory for this recognition process in word learning, thereby leading to greater learning outcomes. This possibly suggests that methods of vocabulary instruction may benefit from reinforcing the orthographic information of a word to establish stronger links to the meaning of the word.

It is important to note that ERPs are beneficial for distinguishing among learning outcomes that may not be detected by behavioral measures. For example, results from the P600 illustrated comprehension skill differences in recognition that were not observable in the behavioral data. First, we take this as encouraging evidence that cognitive neuroscience methods, such as ERPs, may shed light on issues related to learning, such as vocabulary instruction. Second, these results may suggest that in direct instruction of vocabulary, less-skilled readers may not be able to make use of the information that is available to them to establish high quality word representations. Specifically, from our ERP findings we conclude that less-skilled readers in our sample are deficient in establishing quality representations as their high-skilled reader counterparts. As evidenced from our P600 results, less-skilled readers are not recognizing that they have experienced a word before, which may impede future learning when this word is encountered again. Further, they struggle in establishing links from a word's phonology to its meaning. These deficiencies were further illustrated in the N400 data; failure to establish strong representations from orthography to meaning and especially from phonology to meaning hindered the less-skilled readers' ability to make accurate semantic decisions. Thus, this suggests that for the less-skilled readers, phonological information alone that is linked to meaning is insufficient information in a vocabulary instruction environment. The data suggest that these particular readers will benefit more from orthographic information that is linked to meaning. Reinforcing the word's orthography may be especially a good strategy for learning in these readers to ensure that they are able recognize the word in subsequent encounters, and thus, facilitating the process of incremental learning.

Appendix A

Learned rare word	Definition	Sentence	Meaning probe
Abeyant	Temporarily inactive	The plan was **abeyant** until we could get further funding	Suspended
Afreet	A powerful evil spirit or gigantic monstrous demon in Arabic mythology	To protect herself from the evil **afreet**, she wore a garlic necklace around her neck	Demon
Agog	In eager desire; eager highly excited	She explored the deck of the ship, **agog** with excitement for her first trip	Excited
Alate	Having winglike extensions or parts; winged	The **alate** seeds of the maple floated through the forest on the wind	Winged
Ambry	In churches, a kind of closet, niche, cupboard or locker for utensils, vestments, etc.	After the church service, the priest put the silver chalice back in the **ambry**	Cupboard
Arcuate	Having the form of a bow; curved	The **arcuate** arteries are small curved branches of arteries supplying the brain with fresh blood	Curved
Assuasive	Soothing; calming	His **assuasive** remarks really helped the family in their time of need	Soothing
Badinage	Playful teasing; banter	The playful **badinage** of her coworkers made her office a fun place to work	Teasing
Bandy	To toss or throw back and forth (especially words)	We **bandied** many words about when we tried to come up with a good title	Volley
Bawdy	Humorously coarse; risqué. Vulgar, lewd	The men sat around smoking and telling **bawdy** stories about their youths	Rude
Beleaguer	To harass; annoy persistently	The other children **beleaguered** the boy because of his lisp	Bother
Bibulous	Of, pertaining to, marked by, or given to the consumption of alcoholic drink	After each weekend, the **bibulous** can be found hungover, holding their heads and wearing their sunglasses	Alcoholic
Bivouac	An encampment for the night usually without tents or covering	Our troops retreated for the night and went into the **bivouac**	Camp
Blandish	To flatter with kind words or affectionate actions	She used her ability to **blandish** to her advantage anytime she wanted a raise at her job	Coax
Cabochon	A highly polished, rounded stone	The sapphire was cut like a **cabochon** for the ring	Gem
Cerulean	Azure; sky-blue	The **cerulean** skies were mesmerizing to Pittsburghers who see gray, cloudy skies most of the year	Blue

(*continued on next page*)

Appendix A (*continued*)

Learned rare word	Definition	Sentence	Meaning probe
Chafferer	A vendor who enjoys talking while making a sale; a bargainer	Most street vendors are **chafferers** by nature because they like to sell their merchandise at the best price	Bargainer
Chaparral	A dense thicket of shrubs and small trees	Once she lost her necklace in the **chaparral,** she knew she would never see it again because it was too dense to look through	Forest
Clement	Inclined to be lenient or merciful; mild	The weather was particularly **clement,** so it was the perfect day for a walk in the park	Mild
Cygnet	A young swan	The young **cygnet** was swimming closely next to his swan mother	Swan
Dandle	To move (a small child) up and down on the knees or in the arms in playful way; to pamper or pet	It is an old wives' tail that if you don't **dandle** your baby on your lap, he or she will get fat	Bounce
Diptych	An ancient writing tablet	Ancient Greeks used **diptychs** to practice writing	Tablet
Ersatz	Imitation, fake, artificial	She wrote a letter full or **ersatz** sympathy to the coworker she never liked, who had just been fired	Synthetic
Estival	Of, relating to, or appearing in summer	Jenny's family always liked to spend some time **estival** at the Hampton's	Summer
Famulus	A private secretary or other close attendant	In the late 1800's, it was typical for rich families to have a **famulus** working in their home	Attendant
Flivver	An automobile, especially one that is small, inexpensive, and old	John decided to buy a new car after his old **flivver** broke down on the highway	Car
Folderol	Foolishness, nonsense	Her silly **folderol** comments discredited her ability to serve as governor	Nonsense
Fossick	To search for gold in abandoned claims or to rummage around for anything valuable	The homeless tend to **fossick** through trash, hoping to find something they can use	Rummage
Frisson	A shudder of excitement	As we descended from the pinnacle of the rollercoaster track, we experienced a short **frisson** of excitement	Shudder
Garboil	Confusion; uproar	To avoid **garboil**, the twins never wore the same clothes	Confusion
Girandole	An ornate candle holder; often with a mirror	She said that the antique **girandole** hanging on the wall is also known as a "mirror chandelier"	Candles
Gloaming	Twilight; dusk; the fall of the evening	He arrived at the village station on a wintry evening, when the **gloaming** was punctuated by the cheery household lamps	Dusk
Glossal	Of or relating to the tongue	His **glossal** nerve had been severed, leaving him with no feeling in his tongue	Tongue
Gravamen	The most important part of a complaint	The **gravamen** found at the crime scene was sure to convict him of the murder	Complaint
Grimalkin	A cat, especially an old female cat; an old woman considered to be ill-tempered	The irritable **grimalkin** would not play with the other cats because she was too old to run around the house	Cats
Hebetude	Mental dullness or sluggishness	Some say that too much television is leading us toward a nationwide **hebetude**	Dullness
Heptarchy	A government by seven persons; also, a country under seven rulers	The local government rid itself of **heptarchy** because of the many disagreements between the government officials	Seven
Hinny	The hybrid offspring of a male horse and female donkey	At first glance she thought the **hinny** looked like a donkey, but on closer inspection she decided the animal is more subtly like a horse	Animal
Hubris	Overbearing pride or presumption	With dizzying **hubris,** Shelley elevated the purpose of the poet over that of priest and statesman	Pride
Ibex	Wild goat of mountain areas of Eurasia and N. Africa having large backward curving horns	The **ibex** wandered freely throughout the African mountains	Goat
Illation	A conclusion, a deduction, or an inference	Faulty deductions or unimportant **illations** form a false image of things	Inference
Intarsia	A decorative inlaid pattern in a surface, especially	To construct an **instarsia,** outline drawings are	Mosaic

Appendix A (*continued*)

Learned rare word	Definition	Sentence	Meaning probe
	a mosaic worked in wood	used as templates for cutting the many pieces of wood	
Jubilee	A season or an occasion of joyful celebration	Jessica was enjoying the celebrations of the annual spring **jubilee** with her family	Anniversary
Junto	A small, usually secret, group joining for a common purpose	Chris joined the **junto** in hopes of meeting other Star Trek aficionados	Group
Kilderkin	A cask; a small barrel; an old liquid measure containing 18 English beer gallons, or nearly 22 gallons	The tavern kept its beer supply in **kilderkins** behind the bar	Barrel
Kittle	Touchy; unpredictable	Because of her **kittle** personality, her friends never knew what to expect from her when they spent time together	Unpredictable
Kyphosis	Abnormal rearward curvature of the spine, resulting in protuberance of the upper back; hunchback	The hunchback of Notre Dame had severe **kyphosis**	Humpback
Lacrimal	Of or relating to tears	Tears are formed in the **lacrimal** gland, which is located under the upper eyelid	Crying
Lambaste	To give a thrashing to; to beat severely; to scold sharply; to attack verbally; to berate	The politician spent most of his campaign money **lambasting** his opponent rather than discussing the issues	Criticize
Legate	An ambassador or envoy	President Bush sent a **legate** to the Middle East to represent the United States at the peace talks	Representative
Leister	A three-pronged spear used in fishing	Native Americans used **leisters** to catch their fish long ago	Spear
Maculate	To spot, blemish, or pollute	She accidentally dropped her white scarf in the muddy puddle; the **maculate** was hard to get out	Stain
Monish	To warn	His mother **monished** him of the consequences for driving without a seat belt	Warn
Myxoid	Containing mucous; mucoid	She had a **myxoid** cyst at the end of her finger	Mucous
Napiform	Turnip-shaped; large and round in the upper part, and very slender below	The plant had a **napiform** root, and the large upper part could be eaten	Turnip
Natant	Floating or swimming in water	She wanted to get some **natant** plants for the pond she dug in her yard	Floating
Nebbish	A weak-willed, timid, or ineffectual person	He is a **nebbish** person who might be played effectively by Woody Allen	Timid
Oblation	The act of offering something, such as worship or thanks, especially to a diety	The priest reminded everyone to offer their **oblation** throughout the week, not just on Sundays	Worship
Onus	A burden; an obligation; a difficult or disagreeable necessity	The **onus** was on the prosecution to prove the man had gotten sick because of his work environment	Burden
Paranymph	The bridesmaid conducting the bride to the bridegroom	The bride has one chief **paranymph** that helps her out with the wedding plans; she is also known as the maid of honor	Bridesmaid
Peruke	A wig, especially one worn by men in the 17th and 18th centuries	In addition to wearing a **peruke,** Lord Wadsworth sold them to other men wanting long, flowing hair	Wig
Piaffe	A cadenced trot executed by the horse in one spot	To win the equestrian match, John trained his horse to do a fancy **piaffe** in front of the judges	Trot
Pillory	A wooden framework on a post, with holes for the head and hands, in which offenders were formerly locked to be exposed to public scorn as punishment; to expose to ridicule and abuse	As a punishment for adultery, she was locked in a wooden **pillory** for a week at the center of town	Punishment
Pintle	A hook or bolt on the rear of a gun or trailer; the pin on which a gun carriage revolves; the pin on which a rudder turns; a pin or a bolt on which another part pivots; the one that holds a hinge together	The rusty **pintles** must be replaced soon before they a towing vehicle for attaching snap, causing the door to fall off the hinges	Pin
Pluvious	Characterized by heavy rainfall; rainy	The **pluvious** weather lasted for days, causing a devastating flood throughout the small town	Rain
Prescient	Having foresight or knowledge of what will	His **prescient** that the Buccaneers would win the	Intuition

(*continued on next page*)

Appendix A (*continued*)

Learned rare word	Definition	Sentence	Meaning probe
	happen	Superbowl was correct	
Quahog	An edible clam of the Atlantic coast of N. America, having a hard rounded shell	The restaurant used only the freshest **quahogs** to make their famous chowder	Clam
Quisling	A traitor who serves as the puppet of the enemy occupying his or her country	Hitler's plan was to use Cuesta as his Mexican **quisling**	Traitor
Ramous	Of or resembling branches	Some cancers grow in a **ramous** fashion, resembling tree limbs	Branching
Refluent	Flowing back, ebbing	The tide is **refluent,** so we'll soon be able to walk further down the beach	Returning
Rivage	A coast, shore, or bank	She liked to wander along the green **ravage** in the spring, listening to the river	Seashore
Roorback	A false or slanderous story used for political advantage	Politicians use **roorbacks** to defame the name of their opponents.	Slander
Salver	A tray for serving food or drinks	She put the drinks on the **salver** and headed out of the kitchen to serve them to her guests.	Tray
Schism	A separation or division into factions	The **schism** between East and West Germany ended when the Berlin Wall came down.	Split
Sibilant	Of, characterized by, or producing a hissing sound like that of (s) or (sh)	The poet used a lot of **sibilant** consonants in his poem about snakes.	Hiss
Solfeggio	A singing exercise using the syllables: do, re, me, fa, so la, ti	Singers use the **solfeggio** exercise to warm up their vocal cords before they start singing.	Sing
Solleret	A flexible steel shoe forming part of a medieval suit of armor.	During battle, the knights kept their feet protected with steel **sollerets.**	Armor
Temblor	An earthquake	The **temblor** caused destruction throughout the city, and some could feel the tremor from miles outside of town.	Earthquake
Ulster	A loose, long overcoat made of heavy, rugged fabric and often belted	It is appropriate to wear a dark **ulster** over a dinner jacket or tuxedo.	Coat
Ultima	Most remote; furthest; final; last; the last syllable of a word	They're traveling across Europe by train, and are planning on an **ultima** stop in Moscow.	Final
Unco	So unusual as to be surprising; uncanny; Extraordinary	He was a great traveling companion because of his **unco** sense of direction	Strange
Uranic	Of or relating to the heavens; celestial	It is an **uranic** principle that performing evil acts will bring punishment to the evil-doer.	Heaven
Ursine	Of or characteristic of bears or a bear	Because of its **ursine** appearance, the great panda has been identified with the bears; actually, it is closely related to the raccoon.	Bear
Urticant	Causing itching or stinging	The sea anemone tentacles can be used to attack because of their **urticant** properties.	Itch
Venatic	Of or relating to hunting	John always liked **venatic** sports such as deer and duck hunting.	Hunt
Venial	Forgivable; excusable; pardonable	Eating meat on a Friday is a **venial** sin, but murder is a mortal sin.	Forgiven
Vesicate	To blister or become blistered	Her feet **vesticated** after wearing ill-fitting high-heeled shoes all day.	Blister
Vilipend	To treat something as if it has little value; to express a low opinion of	He thought I was **vilipending** his effort, but in actuality I appreciated his work very much.	Belittle
Virgulate	Shaped like a small rod	The **virgulate** shape of the branch made a great makeshift stake that she substituted for the lost tent stake	Rod
Vulpine	Of, resembling, or characteristic of a fox; cunning; clever	The sly **vulpine** lurked behind the trees preparing to attack his prey.	Fox
Wheedle	To entice by soft words; to cajole; to flatter; to coax; to gain or get by with flattery or guile	The school was always trying to **wheedle** contributions from the parents by telling them how bright their children were	Persuade
Wyvern	a two-legged dragon having wings and a barbed tail	The brave knight fought off the ferocious **wyvern** and saved the kingdom	Dragon
Xanthous	Yellow; having light brown or yellowish skin	The **xanthous** coloring of her skin suggested she contracted j aundice	Yellowish

Appendix A (*continued*)

Learned rare word	Definition	Sentence	Meaning probe
Xerosis	Abnormal dryness, especially of the skin and eyes	She always had horrible **xerosis** in the winter, and bought practically every bottle of lotion in the store	Dry
Xylograph	An engraving on wood	She bought some beautiful **xylographs** to hang on her office wall	Engrave
Xyloid	Resembling wood; having the nature of wood	Wormwood is neither worm-shaped or **xyloid**	Wooden
Yagi	A sharply directional antenna	John adjusted the **yagi** on his television because he wanted better reception of the football game	Antenna
Yashmak	A veil worn by Muslim women to cover their face in public.	In some middle eastern countries, women who do not wear their **yashmaks** in public can be arrested by the police	Veil
Yenta	a person, especially a woman, who is meddlesome meddlesome or gossipy	The **yenta** is famous for knowing the latest rumors in town	Gossip
Younker	a young person (especially a young man or boy)	Jan has a three year-old **younker** and two older daughters	Child
Yttrium	a silvery metallic element	They learned in Chemistry class that **Yttrium** Oxide has industrial uses for the manufacturing of television sets	Metal
Zecchino	sequin; a spangle often sewn on cloth	Mary was wearing a beautiful dress with golden **zecchinos** that sparkled	Sequin
Zucchetto	a skullcap worn by certain Roman Catholic clerics	One can tell the rank of a Roman Catholic cleric by the color of the **zucchetto** they wear on their heads	Cap
Zydeco	popular music of southern louisiana played by small groups featuring the guitar, the accordion, and a washboard	They could tell the carnival celebration was about to begin when the **zydeco** melodies were heard several street blocks away	Music

Familiar words		Unlearned rare words	
First word	Meaning probe	First word	Meaning probe
Appendage	Limb	Balatron	Buffoon
Arbitrate	Decide	Buccula	Chin
Aspire	Hope	Burghbote	Donation
Avid	Eager	Caitiff	Coward
Bewilder	Puzzle	Chiliada	Thousand
Botany	Plants	Cullion	Rascal
Brochure	Pamphlet	Demiurge	Creator
Enamel	Coating	Dithyramb	Hymn
Goulash	Stew	Eroteme	Symbol
Hobble	Limp	Jactancy	Brag
Horrify	Scare	Jowter	Salesman
Infantile	Immature	Kevel	Hammer
Intellect	Mind	Limacine	Slug
Jettison	Discard	Lohock	Antibiotic
Kayak	Boat	Lorgnette	Glasses
Lacerate	Slash	Martext	Preacher
Mannequin	Model	Muzhik	Peasant
Mimic	Imitate	Novercal	Stepmother
Moccasin	Slipper	Oubliette	Dungeon
Molecule	Particle	Oxter	Armpit
Muggy	Humid	Poulaine	Shoe
Nauseous	Queasy	Oiviut	Wool
Novice	Beginner	Ructation	Burp
Obligate	Commit	Saulie	Mourner
Organic	Natural	Sororate	Marriage
Parcel	Carton	Speculate	Sharpen

(*continued on next page*)

Appendix A (*continued*)

Familiar words		Unlearned rare words	
First word	Meaning probe	First word	Meaning probe
Peril	Danger	Tomalley	Lobster
Prudence	Caution	Tomentose	Hairy
Quantity	Number	Tregetour	Magician
Reluctant	Hesitant	Uloid	Scar
Remedy	Cure	Ustulate	Burnt
Revoke	Withdraw	Uxorcide	Murder
Sentiment	Emotion	Wibrissae	Whiskers
Skittish	Restless	Yataghan	Knife
Slaughter	Kill	Zori	Sandle

References

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (p. 328). Newark, DE: International Reading Association.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., et al. (2007). The English Lexicon Project. *Behavior Research Methods, 39*, 445–459.

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition, 10*, 82–93.

Beck, I. L., McKeown, M. G., & Omanson, R. C. (1987). The effects and uses of diverse vocabulary instructional techniques. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 147–163). Hillsdale, NJ: Erlbaum.

Bentin, S. (1989). Electrophysiological studies of visual word perception, lexical organization, and semantic processing: A tutorial review. *Language and Speech, 32*(2), 205–220.

Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: Time course and scalp distribution. *Journal of Cognitive Neuroscience, 11*, 235–260.

Bolger, D. J., Balass, M., Landen, E., & Perfetti, C. A. (2008). Contextual variation and definitions in learning the meanings of words. *Discourse Processes, 45*(2), 122–159.

Cain, K., Lemmon, K., & Oakhill, J. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge and memory capacity. *Journal of Educational Psychology, 96*, 671–681.

Carver, R. P. (1992). What do standardized tests of reading comprehension measure in terms of efficiency, accuracy, and rate? *Reading Research Quarterly, 27*(4), 346–359.

Chaffin, R. (1997). Associations to unfamiliar words: Learning the meanings of new words. *Memory & Cognition, 25*(2), 203–226.

Creaser, J., Jacobs, M., Zaccaria, L., & Carsello, C. (1970). Effects of shortened time limits on the Nelson–Denny reading test. *Journal of Reading, 14*(3), 167–170.

Cummins, R. P. (1981). Test review: The Nelson–Denny reading test (forms E and F). *Journal of Reading, 25*(1), 54–59.

Curran, T. (1999). The electrophysiology of incidental and intentional retrieval: ERP old/new effect in lexical decision and recognition memory. *Neuropsychologia, 37*, 771–785.

Curtis, M. E. (1987). Vocabulary testing and vocabulary instruction. In M. G. McKeown & M. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 37–51). Hillsdale, NJ: Erlbaum.

Daneman, M. (1984). Individual differences in reading skills. In Rebecca Barr & Michael L. Kamil (Eds.). *Handbook of reading research* (Vol. 2, pp. 512–538). Lawrence Erlbaum Associates, Inc.

Dixon, P., LeFevre, J., & Twilley, L. C. (1988). Word knowledge and working memory as predictors of reading skill. *Journal of Educational Psychology, 80*, 465–472.

Drum, P. A., & Konopak, B. C. (1987). Learning word meanings from written context. In M. G. Margaret & M. E. Curtis (Eds.). *The nature of vocabulary acquisition* (Vol. ix, pp. 73–87). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc..

Everatt, J., & Underwood, G. (1994). Individual differences in reading subprocesses: Relationships between reading ability, lexical access, and eye movement control. *Language and Speech, 37*(3), 283–297.

Ferree, T. C. (2006). Spherical splines and average referencing in scalp electroencephalography. *Brain Topography, 19*(1/2), 43–52.

Finnigan, S., Humphreys, M. S., Dennis, S., & Geffen, G. (2002). ERP 'old/new' effects: Memory strength and decisional factor(s). *Neuropsychologia, 40*, 2288–2304.

Fukkink, R. G., & de Glopper, K. (1998). Effects of instruction in deriving word meanings from context: A meta-analysis. *Review of Educational Research, 68*, 450–469.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*(2), 95–112.

Hart, L. A. (2005). A training study using an artificial orthography: Effects of reading experience, lexical quality, and text comprehension in L1 and L2. Unpublished doctoral dissertation, University of Pittsburgh, Pittsburgh.

Herman, P. A., Anderson, R. C., Pearson, P. D., & Nagy, W. E. (1987). Incidental acquisition of word meaning from expositions with varied text features. *Reading Research Quarterly, 22*, 263–264.

Jenkins, J. R., Stein, M., & Wysocki, K. (1984). Learning vocabulary through reading. *American Educational Research Journal, 21*(4), 767–787.

Key, A. P. F., Molfese, D. L., & Ratajczak, E. D. (2006). ERP indicators of learning in adults. *Developmental Neuropsychology, 29*(2), 379–395.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203–205.

Landi, N. (2005). Behavioral and electrophysiological investigations of semantic processing in skilled and less-skilled comprehenders. Unpublished doctoral dissertation, University of Pittsburgh, Pittsburgh.

McCandliss, B., Posner, M., & Givón, T. (1997). Brain plasticity in learning visual words. *Cognitive Psychology, 33*, 88–110.

McKeown, M. G. (1985). The acquisition of word meaning from context by children of high and low ability. *Reading Research Quarterly, 20*(4), 482–496.

Miller, G. A. (1999). On knowing a word. *Annual Review of Psychology, 50*, 1–19.

Miller, G. A., & Charles, W. G. (1991). Contextual correlates of semantic similarity. *Language and Cognitive Processes, 6*(1), 1–28.

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly, 20*(2), 233–253.

Nelson, M. J., & Denny, E. C. (1973). The Nelson–Denny reading test. Houghton Mifflin Company.

Nelson, J. R., Balass, M., & Perfetti, C. A. (2005). Differences between written and spoken input in learning new words. *Written Language & Literacy, 8*(2), 25–44.

Nelson, J.R., & Halderman, L.K., (2008, October). A principle components analysis of a database of reading measures. Talk presented at the Charles A. Perfetti laboratory meetings.

Nobre, A. C., & McCarthy, G. (1994). Language-related ERPs: Scalp distributions and modulation by word type and semantic priming. *Journal of Cognitive Neuroscience, 6*(3), 233–255.

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*, 554–566.

Patterson, K., & Hodges, J. R. (1992). Deterioration of word meaning: Implications for reading. *Neuropsychologia, 30*(12), 1025–1040.

Perfetti, C. A. (1985). *Reading ability*. New York: Oxford Press.

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357–383.

Perfetti, C. A., & Hart, L. (2001). The lexical basis of comprehension skill. In David. S. Gorfein (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (pp. 67–86). Washington, DC: American Psychological Association.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Vehoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). Amsterdam/Philadelphia: John Benjamins.

Perfetti, C. A., Wlotko, E. W., & Hart, L. A. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(6), 1281–1292.

Raven, J. C. (1960). *Guide to using the standard progressive matrices*. London: Lewis.

Reichle, E. D., & Perfetti, C. A. (2003). Morphology in word identification: A word experience model that accounts for morpheme frequency effects. *Scientific Studies of Reading, 7*, 219–237.

Rugg, M. D., Allan, K., & Birch, C. S. (2000). Electrophysiological evidence for the modulation of retrieval orientation by depth of study processing. *Journal of Cognitive Neuroscience, 12*(4), 664–678.

Scherg, M., Ille, N., Bornfleth, H., & Berg, P. (2002). Advances tools for digital EEG review: Virtual source montages, whole-head mapping, correlation, and phase analysis. *Journal of Clinical Neurophysiology, 19*(2), 91–112.

Sternberg, R. J., & Powell, J. S. (1983). Comprehending verbal comprehension. *American Psychologist, 38*(8), 878–893.

van Daalen-Kapteijns, M., & Elshout-Mohr, M. (1981). The acquisition of word meanings as a cognitive learning process. *Journal of Verbal Learning and Verbal Behavior, 20*, 386–399.

Widling, E. L., & Rugg, M. D. (1997). An event-related potential study of memory for words spoken aloud or heard. *Neuropsychologia, 35*(9), 1185–1195.