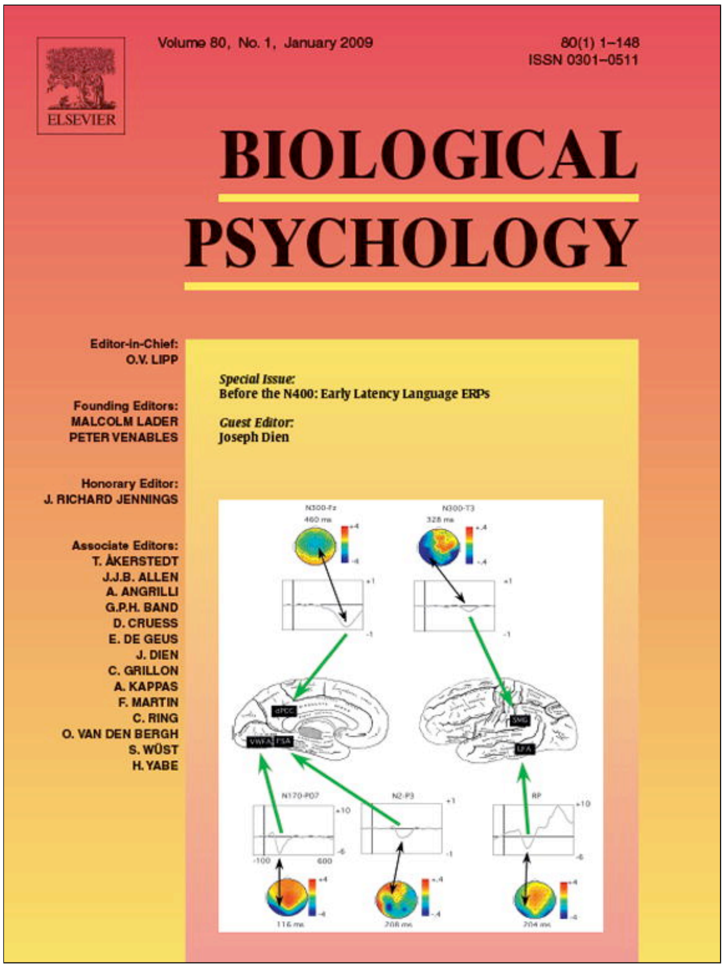

Volume 80, No. 1, January 2009
80(1) 1–148
ISSN 0301-0511
ELSEVIER
BIOLOGICAL
PSYCHOLOGY
Editor-in-Chief:
O.V. LIPP
Founding Editors:
MALCOLM LADER
PETER VENABLES
Honorary Editor:
J. RICHARD JENNINGS
Associate Editors:
T. ÅKERSTEDT
J.J.B. ALLEN
A. ANGRILLI
G.P.H. BAND
D. CRUESS
E. DE GEUS
J. DIEN
C. GRILLON
A. KAPPAS
F. MARTIN
C. RING
O. VAN DEN BERGH
S. WÜST
H. YABE
Special Issue:
Before the N400: Early Latency Language ERPs
Guest Editor:
Joseph Dien

Biological Psychology 80 (2009) 130–147

Contents lists available at ScienceDirect

Biological Psychology

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

ERP measures of partial semantic knowledge: Left temporal indices of skill differences and lexical quality

Gwen A. Frishkoff [a,*], Charles A. Perfetti [a], Chris Westbury [b]

[a] Learning Research & Development Center, University of Pittsburgh, Pittsburgh, PA, United States
[b] University of Alberta, Alberta, Canada

ARTICLE INFO

Article history:
Received 13 November 2007
Accepted 22 April 2008
Available online 13 May 2008

Keywords:
Event-related potentials
Semantic processing
Word learning
Individual differences
Reading

ABSTRACT

This study examines the sensitivity of early event-related potentials (ERPs) to degrees of word semantic knowledge. Participants with strong, average, or weak vocabulary skills made speeded lexical decisions to letter strings. To represent the full spectrum of word knowledge among adult native-English speakers, we used rare words that were orthographically matched with more familiar words and with pseudowords. Since the lexical decision could not reliably be made on the basis of word form, subjects were obliged to use semantic knowledge to perform the task. A d' analysis suggested that high-skilled subjects adopted a more conservative strategy in response to rare versus more familiar words. Moreover, the high-skilled participants showed a trend towards an enhanced "N2c" to rare words, and a similar posterior temporal effect reached significance ~650 ms. Generators for these effects were localized to left temporal cortex. We discuss implications of these results for word learning and for theories of lexical semantic access.

1. Introduction

Although we commonly speak of knowing or not knowing a word, the clear lesson from vocabulary research is that knowledge of a word's meaning is not all-or-none (Dale, 1965; Stahl, 1986). Rather, robust word knowledge typically requires multiple exposures to a word in different contexts (Beck et al., 2002). Probing comprehension after a few exposures may reveal a partial grasp of the semantics — e.g., the fact that *adventure* is associated with some kind of activity and is typically enjoyable — although the comprehender is unable to provide a precise definition, or to use the word in an appropriate context (Brown et al., 2005; Frishkoff et al., in press). The implication is that semantic knowledge is acquired incrementally. Thus, at any time, a learner may have graded and incomplete knowledge of meanings for some words. While this idea has intuitive appeal, there have been remarkably few studies focused on the neurocognitive processes that may be associated with different degrees of semantic knowledge. The idea that word knowledge can be partial has been widespread in the educational literature, but has received little attention in neurocognitive studies of semantics.

The goal of the present experiment was to examine modulation of event-related brain potentials (ERPs) in response to words that are partially known, as compared to words that are very rare, and therefore likely to be novel, and to words that are more familiar. In the word learning literature, words that are partially known have been designated as "frontier words" (Durso and Shore, 1991). Frontier words are familiar in form, but their semantic representation is either incomplete or unstable, or both. Recent work on the neurocognition of word semantic processing has described how frontier words may engage qualitatively different processes when compared with words that are either well known or completely novel (Ince and Christman, 2002).

In the present experiment, three classes of words (low-frequency, rare, and very rare words) were presented in a lexical decision task. Several features of the task design promoted implicit access to word meaning as the means to differentiate words from nonwords. Thus, the emphasis was on indirect and implicit, rather than direct and conscious, semantic access. In keeping with the theme for this special issue, our focus is on early ERP markers of semantic knowledge (peaking <400 ms). Of particular interest are the posterior temporal N2c, the left anterior N3, and the medial frontal MFN components, peaking at approximately 200–250 ms, 300–350 ms, and 350–400 ms, respectively. As detailed below, these patterns have been observed in previous studies of word-level semantic processing. While their spatiotemporal attributes differ across studies, and their functional significance remains to be clarified, these patterns appear to reflect distinct aspects of

* Corresponding author at: Learning Research and Development Center, 3939 O'Hara Street, Office 642, University of Pittsburgh, Pittsburgh, PA 15260, United States. Tel.: +1 412 260 8010.
E-mail address: gwenf@pitt.edu (G.A. Frishkoff).

0301-0511/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.biopsycho.2008.04.017

word recognition. The question for the present study was how these patterns are modulated in response to words that differ in familiarity, particularly when the task involves implicit access to semantic knowledge. To differentiate these components, we applied a Principal Components Analysis and examined both the topography and timing of each principal component. For further insights into the neural bases of these patterns, we modeled anatomical sources using a minimum-norm least squares algorithm (MNLS). Based on previous results, we expected to find activity in left posterior temporal (fusiform), left anterior inferior temporal, and mid-frontal regions. We further expected to find increased activity in left temporal regions for frontier versus familiar words, reflecting the need for additional effort to retrieve lexical and semantic information associated with these words. Given that word knowledge is closely linked to vocabulary and reading comprehension abilities, we also examined skill differences in lexical semantic access. Participants were divided into three skill groups, based on a standardized test of vocabulary knowledge. Analysis of the *d*-prime "beta criterion" for the lexical decision task served as an index of different strategies. This measure was used to interpret differences in the ERP effects for rare words versus low-frequency words and pseudowords among the three groups of participants.

1.1. Degrees of word knowledge: The cognitive status of "frontier" words

Children have been estimated to acquire between 2000 and 3000 words per year (Nagy et al., 1987). Adults likewise continue to add to their vocabulary, though at a more modest rate. This is particularly the case for second- (and third- and fourth-) language learners (cf. Proverbio et al., in press). This ability to learn new words with relative ease, even as adults, contrasts with the well-known limits on adult acquisition of phonological and morpho-syntactic knowledge. It seems that the neural systems for encoding meaning may remain fairly plastic throughout development (cf. Neville et al., 1992), although word learning processes and strategies are likely to develop with maturation and experience.

Given the incremental nature of word learning and the rate of vocabulary growth, particularly in childhood, it is likely that a large number of words are partially known at any given time. The main question for the present study was whether a task that promotes implicit semantic access would evoke the same ERP patterns for rare (or "frontier") words as for more familiar (low-frequency) words, and for words that are completely novel. Our hypothesis was that frontier words would evoke qualitatively different patterns, reflecting recruitment of specific cognitive and neural processes that are implicated in early semantic processing. An alternative prediction to be tested was that all words recruit the same neurocognitive networks, but that strength of activation within these networks would vary with the strength of the memory representation for a particular word, without affecting the overall pattern across these networks. That is, the neural response to frontier words may be quantitatively, but not qualitatively, different from the response to more familiar words.

There are several reasons for expecting to find qualitative differences in processing frontier versus known words. First, word representations are multidimensional, comprising orthographic, phonological, semantic, and morphosyntactic features. Further, a word can be experienced in contexts that draw attention to one or more of these features. It is therefore possible that different word representations vary in the relative strength of these features, leading to different patterns of neural processing. In addition to the multiple dimensions of word form, meaning is itself complex. For example, some theories of semantic processing have emphasized differences between concrete and abstract meanings (Paivio et al., 1988), and event-related potentials have been shown to differ systematically in response to words that are rated as either highly concrete or highly abstract, consistent with dual "routes" for processing concrete versus abstract meanings (Kounios and Holcomb, 1994; but cf. Pexman et al., 2007 for a different view). The complexity of words makes it easy to see how word representations can vary in their "quality" or depth (Perfetti and Hart, 2001, 2002).

In summary, the cognitive and neural processes that give rise to semantic knowledge are thought to remain plastic throughout development. This may explain why word meanings can be acquired at any age (Tucker, 2002). The plasticity of semantic knowledge is also consistent with the observation that word learning is correlated with frequency of exposure to a word. In other words, word learning is experience-dependent: the words that an individual knows well are likely to be ones that he or she has frequently encountered (Reichle and Perfetti, 2003). At the same time, because words and their semantic associations are complex, there are good reasons to expect that partially known words may evoke different cognitive and neural processes than well-known words, particularly in tasks that involve access to meaning.

1.2. Skill differences in lexical processing: The Lexical Quality Hypothesis

The words that a person knows depend not only on frequency of exposure, but also on that person's vocabulary and other linguistic skills (Cain et al., 2003). For example, high-skilled readers learn new word meanings from written texts more readily than low-skilled readers (Swanborn and De Glopper, 2002; McKeown, 1985). In examining neurocognitive responses to partially known words, it may therefore be important to consider the role of vocabulary skills and other language abilities that may be related to differences in word recognition.

Skill differences may reflect a number of underlying factors, including text comprehension abilities, working memory skills, and vocabulary size. To explore the structure of word knowledge and its relation to various cognitive and linguistic skills, Perfetti and Hart (2001, 2002) developed a large battery of adult reading and language tests. They report a factor analysis of performance by 445 adults on a subset of tests of this battery. The factor structures that emerged suggest a more coherent lexical knowledge structure for skilled readers than less-skilled comprehenders. Vocabulary knowledge was highly related to knowledge of word form (spelling and pronunciation) for skilled readers, loading on a single factor. But for less-skilled comprehenders, this relationship was weaker, with vocabulary knowledge loading on a factor separated from form knowledge. These results point to qualitative differences in lexical representations for strong and weak readers, which may reflect a stronger coherence among different dimensions of word knowledge for stronger readers. Perfetti and Hart refer to this idea as the Lexical Quality Hypothesis. This hypothesis may provide one account of how skilled and less-skilled readers may differ in their representation of word knowledge. While word knowledge has multiple dimensions (phonology, orthography, semantics), it may be the case that these dimensions must be cognitively bound to form a complete, integrated lexical representation. If so, then the results from Perfetti and Hart suggest that less-skilled comprehenders lack sufficient experience in reading to have developed the necessary links between form-based and semantic skills that can help to establish new form–meaning relationships.

In the present study, we measured individual differences in reading and language abilities using a battery of lexical knowledge and reading comprehension tests that was developed previously in our lab (Perfetti and Hart, 2001, 2002). Using subject scores on the Nelson–Denny vocabulary test, we classified participants into three skill groups. Given the prior research suggesting skill differences in word learning and lexical processing, we hypothesized that groups would differ specifically in processing of "frontier" words, and that these differences would be reflected in ERP measures of early semantic processing.

1.3. ERP studies of lexical semantic processing

Previous ERP studies have identified several components that appear to be sensitive to lexical semantic manipulations. We restrict our brief review to those studies that have used high channel counts (50 or more channels) and have subjected ERP data to careful spatiotemporal analysis, to distinguish processes that unfold within ~200–400 ms after onset of a visual word stimulus.

A seminal study by Nobre and McCarthy (1994) reported the first high-density (50-channel) ERP recordings to words in various conditions, including a semantic priming manipulation. They identified at least two distinct patterns within the typical "N400" window (~250–500 ms) that were responsive to priming — a left temporal "N330" (N3) and a centroparietal "N364" (N4). The N3 and N4 components are clearly separable in this study: while the N3 is larger (more negative) for semantically primed words, the N4 is smaller (more positive). In addition, a medial frontal component is coincident with the N3 at 330 ms and may be overlapping with the parietal N4 by 360 ms. This pattern appears similar to what we have called the medial frontal negativity (MFN) in prior work (Frishkoff, 2007; Frishkoff et al., 2004; Tucker et al., 2003).

A study by Bentin et al. (1999) contrasted ERP responses to the same linguistic (word and nonword) stimuli under different task conditions (phonological, orthographic, and semantic instructions). The semantic categorization task alone evoked a medial frontal "N450," which began around 300 or 350 ms was larger to pseudowords than to words. This pattern appears similar to the MFN, although its peak latency is a little later than that of the typical MFN (Frishkoff, 2007). In addition, there is a left anterior temporal negativity that appears to peak around 300 ms (the "N320"), similar to an N3. However, in contrast with Nobre and McCarthy (1994), Bentin et al. (1999) report that the N3 is larger for pronounceable pseudowords and words as compared with nonwords in the phonological task, but do not report any differences for this component in the semantic task.

A 64-channel ERP study by Hill et al. (2002) examined differences in lexical priming effects at short (150 ms) and long (700 ms) prime-target stimulus onset asynchronies (SOAs). In addition to the typical P300/N400 priming effects, they report a mid-frontal negativity peaking at around 310 ms, which is more prominent at the short SOA.

Together, these three studies point to the existence of a mid-frontal negativity that peaks around 300–350 ms and may or may not be distinct from the left antero-temporal N3 and from the typical centroparietal N4 that has observed in most studies of semantic processing over the past two decades (see Kutas and Federmeier, 2000 for a review). In a recent study, Frishkoff (2007) applied spatiotemporal source localization to the MFN effect and found that it peaks earlier than the N400/P300 effect and that its variance is mostly accounted for by a source in dorsal anterior cingulate. By contrast, the N400/P300 effect was captured by bilateral sources in medial temporal areas, consistent with other reports (Silva-Pereyra et al., 2003; Rossell et al., 2003; Van Petten and Luka, 2006).

From these results it appears that the MFN is distinct from the N3 and P3/N4 semantic effects. Hill et al. (2002) speculated that this mid-frontal pattern reflects more automatic semantic priming (cf. Neville et al., 1992). It is also possible that this mid-frontal component is related to the mid-frontal "fN400" that has been observed in studies of familiarity-based memory (Curran, 2000; Rugg and Curran, 2007). This leaves open the possibility that the MFN reflects a domain-general process that is activated during lexical semantic retrieval, but is not specific to semantic processing.

In contrast with the mid-frontal MFN, the N2 and N3 components are both strongly left lateralized, suggesting these patterns may be good candidates for early markers of lexical semantic access (Frishkoff, 2007; Frishkoff et al., 2004; Dien et al., 2003). The N3 effect was observed in a greater negativity for primed than less primed words at around 250–300 ms, and was accounted for by sources in left anterior inferior temporal cortex (similar results in Dien et al., 2003).

The left posterior temporal N2 component, or "RP" (recognition potential) has been described in detail by Rudell (1992, 1999) and by Hinojosa and co-workers (Hinojosa et al., 2001; Martin-Loeches et al., 2004). Martin-Loeches et al. (2004) were the first to present clear evidence that this left occipito-temporal negativity might be sensitive to lexical semantic processing: the N2/RP was larger in response to words belong to specific (target) semantic category, in contrast with nontargets. The possibility remains that this early component is not directly implicated in semantic access, but is instead related to orthographic processing that is enhanced through feedback from higher (e.g., frontal or anterior temporal) processes that are themselves related to semantic access. Regardless of its precise role in semantic processing, the N2/RP appears to be the earliest ERP marker of semantic processing that has been consistently described in prior studies.

1.4. Study design and hypotheses

To probe different "degrees" of word knowledge, our experiment made use of the well-known Lexical Decision Task (LDT), with modifications that were designed to capture graded differences in word-level semantic knowledge. First, to represent the full spectrum of word knowledge among adult, native-English speakers, we used *rare* and *very rare words* (such as "limpid" and "nutant," respectively) that were orthographically matched with low-frequency words (such as "abject"). Real words in turn were orthographically matched with nonwords such as "pefitant." Since the lexical decision could not reliably be made on the basis of word form (spelling), subjects were obliged to access their knowledge of word meanings to perform the task (Binder et al., 2003). Second, in addition to response time (RT), we computed *d*-prime (d') measures of word–nonword discrimination to determine how accurately subjects were able to recognize the rare words. D' measures correct for response bias (e.g., individual subject biases towards "word" or "nonword" responses) and are therefore more sensitive than raw accuracy measures. We also collected confidence ratings on each trial to provide a graded measure of familiarity with the meaning of each stimulus. Third, subjects were tested on their knowledge of word meanings on about 12% of the trials (see Section 2). Fourth, task instructs emphasized accuracy over speed, which may increase reliance on word meaning (Binder et al., 2003). Together, these features of the task were designed to encourage reliance on semantic access. At the same time, this access is likely to be more automatic or implicit than in tasks that explicitly require semantic judgments. We therefore expected to find patterns that are related to early automatic semantic access.

The use of event-related potentials (ERPs) is particularly well suited for examining multiple dimensions of word meaning. While speed and accuracy may suggest quantitative differences in performance as subjects are confronted with words they know very well, only somewhat, or not at all, ERP measures may point to specific subcomponents (time windows and patterns of brain activity) that are associated with different stages or different dimensions of word-level semantic processing. As discussed in the previous section, several high-density ERP studies have described multiple ERP components in semantic processing, including early left-hemisphere components (the N2 and N3 patterns), as well as midline components peaking slightly later (the MFN and N400).

To examine skill differences in processing frontier words, we classified participants as having strong, average, or weak vocabulary skills on the bases of a standardized test of vocabulary knowledge (the Nelson–Denny vocabulary test; Brown et al., 1993). We predicted that participants with strong vocabulary skills would show greater word–nonword discrimination in each category (low-frequency, rare, and very rare) and that they would also display greater fluency in lexical processing (that is, faster RTs) and greater self-rated confidence in their familiarity with the meanings of words. We further predicted that high-skilled participants would show more pronounced differentiation in their behavioral and ERP responses to words that were known well versus words that were somewhat familiar (i.e., rare words that they correctly recognized as words) and words that were only vaguely familiar or completely novel (i.e., rare words that they misclassified as nonwords). If confirmed, these findings would point to qualitative, as well as quantitative, differences in word processing among individuals with different vocabulary (and related linguistic) skills.

2. Methods

2.1. Subjects

Participants were 34 monolingual native-English speakers who were right-handed, with normal or corrected-to-normal vision. The average age was 21 (S.D. = 3.5). Nine subjects (26%) were male. Subjects were recruited from a reading and language "prescreening" (see Section 2.2 below for details). Either payment ($20 per hour) or academic course credit, or a combination, was given in exchange for subject participation.

Based on their Nelson–Denny vocabulary scores (see below), subjects were divided into three groups: subjects with *strong vocabulary skills* ($n = 13$) scored in the top third of the distribution, subjects with *average vocabulary skills* ($n = 10$) scored in the central third of the distribution, and subjects with *weak vocabulary skills* ($n = 11$) scored in the bottom third. No attempt was made to control for other skill attributes (see Table 1 for cognitive and reading skill profiles).

2.2. Prescreening: The lexical knowledge battery

Prescreening participants completed a battery of language tests developed and administered within the Reading and Language Lab. Subjects were recruited through introductory courses in Psychology at the University of Pittsburgh, and by posting ads in the campus newspaper. Either payment ($7 per hour) or academic course credit, or a combination, was given in exchange for subject participation.

The prescreening tests include measures of phonological and orthographic skills and knowledge, vocabulary, reading comprehension, reading experience, and nonverbal intelligence. Seven tests in all were administered:

- *Phonological Awareness*: In-house test of several phonological subskills (see Perfetti and Hart, 2001)
- *Spelling*: Checklist that includes correctly and incorrectly spelled words (see Perfetti and Hart, 2001)
- *Decoding*: Checklist that includes pseudohomophones — i.e., real words that are incorrectly spelled — and pseudowords that do not sound like real English words (see Perfetti and Hart, 2001)
- *ND comp*: The Nelson–Denny test of reading comprehension (Brown et al., 1993)
- *ND vocab*: The Nelson–Denny test of vocabulary knowledge (Brown et al., 1993)
- *Author Checklist*: The Author Recognition Test (Stanovich and West, 1989)
- *Raven's Matrices*: An abbreviated version of the Advanced Raven's Matrices test (Raven, 1960)

Table 1 below shows the mean scores on the prescreening tests for each subgroup of subjects who participated in the ERP task. Three of the tests are speeded (ND vocabulary and comprehension, and Raven's matrices); therefore, composite scores — reflecting both speed and accuracy — are shown for these tests. Phonological awareness test scores are reported in percent-accuracy. The remaining three tasks — Author Recognition, Spelling, and Decoding —comprised checklists of "target" items (names of real authors, correctly spelled words, and true pseudohomophones) and false "lures" (false author names, misspelled words, and nonwords that were not pseudohomophones). Given the composition of these tasks, we report performance as d' scores (that is, target versus lure discrimination).

Univariate analyses of variance were performed to test for group differences on each test measure. We adopted a p-value of .007 to correct for multiple comparisons. By this criterion, group differences were found on three of the six measures (not including vocabulary, which was highly significant but trivial, since groups were defined on this variable). There was a significant difference on the test of comprehension, $F(2, 33) = 15.33$, $p < .001$. Post-hoc tests confirmed that the strong vocabulary group had higher mean comprehension scores than the average group, and the average group outperformed the weak vocabulary group. Similarly, the groups differed on Author Recognition (a rough index of amount of reading experience) — $F(2, 33) = 10.60$, $p < .001$ — and on the Spelling Test — $F(2, 33) = 8.73$, $p = .001$.

There were no differences between subgroups on our measures of nonverbal intelligence (Raven's matrices), or on measures of phonological awareness or decoding.

2.3. ERP stimuli

Phase I stimuli consisted of 350 words and word-like stimuli. Letters were lower-case Geneva font black, 26 dpi, presented foveally on a white screen. Letter strings subtended a visual angle between 2.5° and 3°. All real-word stimuli were adjectives, 4–8 letters in length (mean = 7.3 letters).

Rare words were selected as follows. An initial list of over 15,000 rare English words were taken from the online database, "International House of Logorrhea" (Chrisomalis, 2005). A short list of ~350 words was extracted from this longer list. Rare words from the larger list were selected if they met the following criteria:

- They were 4–8 letters in length
- They were not foreign or recently borrowed words (e.g., "cabriolet," "fouter" — Fr.)
- Their meanings were not narrow or technical (e.g., "ablaut," "bacillicide")
- They were not composed of word parts (stems, morphemes) that were semantically transparent or were likely to prime familiar words that share the same root (e.g., "galactophagist" — milk drinker).

The goal was to select stimuli that would meet our criteria for either rare or very rare words. *Rare words* were defined a priori as real-word stimuli that were likely to be recognized by only some of the participants, and if recognized would be likely to have less established semantic representations. *Very rare words* were expected to be recognized by few if any of the participants. Importantly, the orthographic form of these words was matched with the familiar word forms. In this way, we aimed to select words that subjects would not know (or would not know well), *not* by virtue of the stimulus meanings or representations themselves, but merely by accident of history.

To anchor our analysis of responses to rare and very rare words, we selected low-frequency, familiar words and pseudowords that were matched with the rare and very rare words, item-by-item, for length and orthographic set size.

Table 1
Vocabulary subgroup scores on reading and language battery

Vocab group	ND vocab[a] (compos.)	ND comp (compos.)	Author checklist (d')	Spelling (d')	Decoding (d')	Phonological awareness (% accuracy)	Raven's matrices (compos.)
Strong ($n = 13$)	68.06 (7.55)	25.94 (4.71)	1.78 (.58)	2.23 (.36)	2.43 (.74)	.78 (.25)	.66 (.21)
Average ($n = 10$)	43.84 (9.71)	20.88 (6.63)	1.26 (.38)	1.90 (.26)	2.03 (.40)	.80 (.68)	.65 (.21)
Weak ($n = 11$)	15.45 (5.61)	10.47 (6.35)	.85 (.39)	1.48 (.33)	1.77 (.81)	.68 (.13)	.60 (.21)

Standard deviations for each measure are given in parentheses.
[a] Subgroups defined on this criterion.

A series of pairwise comparisons revealed that the rare, very rare, and familiar words differed in mean frequency (three different measures of written word frequency were examined: Francis–Kucera, Thorndike, and Celex measures). All stimulus categories were matched on length and orthographic frequency. Orthographic neighborhood is a rather coarse-grained measure of orthographic familiarity. We also examined average bigram frequency of stimuli in each category (cf. Westbury and Buchanan, 2002). Our main concern was that the pseudowords might contain letter patterns that are less common, distinguishing them from real words at a sublexical level (and perhaps explaining why they have not been adopted as real word forms in English). If subjects were attuned to this, even unconsciously, this would confound our analysis of rare word versus pseudoword representations. Interestingly, the opposite was true: the pseudoword stimuli selected for this study had higher average bigram frequency scores than the other three (real-word) categories. Low-frequency, rare, and very rare words did not differ in bigram frequency. If anything, this difference would decrease the likelihood of observing differences in real word versus pseudoword processing. Conversely, we can be fairly confident that any observed effects are not an artifact of differences in sublexical frequencies.

2.4. Experiment procedure

The experiment session began with a block of 20 practice trials. The main task consisted of 5 blocks of 360 trials each. Each trial began with the presentation of a rectangular outline, which cued the subject to initiate the experimental sequence when ready by pressing both the '1' and '2' keys simultaneously using their right middle and index fingers. Subjects were informed that it was not important to press both buttons in synchrony, but that they should press both buttons to begin each trial. This procedure was designed to avoid differential priming of either the "yes" or "no" response prior to onset of the target stimulus.

Once the subject initiated a trial, a central fixation ("+" symbol) appeared for 500 ms (Fig. 1). Next, a letter string was presented for 2500 ms (fixed duration, independent of subject response time). The task was to decide whether the stimulus constituted a real English word, or not. They were asked "to respond as quickly as possible *without making errors*." (cf. Binder et al., 2003). Trials were randomly presented (order determined separately for each subject). "Yes" and "no" (right index and middle) keys were counterbalanced across subjects. Response time and accuracy were recorded simultaneously with the EEG measures.

After the target disappeared from the screen, one or more questions appeared. If the subject responded that the stimulus was a real word, she was asked to indicate how certain she was that she knew the meaning of the word (1–4 scale, 1 = not at all confident). If she responded '1' ("not at all confident"), she was asked how certain she was that she had seen or heard the word previously. If she responded that the stimulus was not a real word, she was asked to indicate her level of confidence. Subjects were encouraged to take their time when responding to the questions and to provide candid answers.

On ~12% (50) of the trials, the stimulus was a low-frequency, familiar word. If the subject correctly responded "yes" (it is a real word), after she completed the first confidence judgment, she was asked to respond to a multiple-choice question of the form, "Which of the following words is closest in meaning to [the target word]"? The multiple-choice questions were reasonably challenging. The objectives were (1) to provide an additional measure of vocabulary knowledge that could be correlated with offline measures (e.g., Nelson–Denny vocabulary scores), and (2) to encourage subjects to be thoughtful and candid when assessing their knowledge of words and word meanings.

After the subject responded to the final question, the brightened box reappeared, cueing the subject that the trial was complete. During the intertrial interval, subjects were permitted to blink, scratch, or adjust their position, after which they were asked to count slowly to 10 to allow the amplifier to recover before using the response keys to advance to the next trial.

2.5. ERP data acquisition and preprocessing

ERPs were recorded using a 128-channel electrode array, with vertex recording reference (Electrical Geodesics, Inc.). Data were sampled at a rate of 500 per second and were amplified with a .01-Hz highpass filter (time constant ~10 s). Trials were edited online for eye blinks and head movements. The remaining trials were segmented into 1400 ms epochs, starting 500 ms before onset of the target word. Segmented data were averaged across trials (within subjects and within condition) and digitally filtered with a 30-Hz lowpass filter. After further channel and subject exclusion, bad (excluded) channels were interpolated. The data were re-referenced to the average of the recording sites, and were baseline corrected (200 ms prior to target stimulus).

2.6. ERP component analysis & source localization

Following Dien and Frishkoff (2005) and Dien et al. (2005), ERP data were subjected to a temporal Principal Components Analysis (PCA). Dien's PCA Matlab Toolbox (v 1.093) was used to implement the analysis (Dien, 2004). After factor extraction and retention (10 factors with eigenvalues >1), we applied Promax rotation (kappa = 4). Three of these factors were outside the time interval of interest (80–700 ms) and were excluded from the analysis. The time course and topography of six of the remaining seven factors are shown in Fig. 7 (factor F550 was omitted from the graphic due to space limitations). Factor scores were averaged over scalp regions of interest (i.e., channels sets) to examine condition effects for each ERP component (see Appendix A for channel groups).

To model the anatomical source activity for each component, we used a minimum-norm least squares solution, as implemented in GeoSource software (Electrical Geodesics, Inc; www.egi.com). The input to the MNLS were PCA factor scores for each component of interest. A finite difference model was specified, based on cortical segmentation of the Montreal Neurologic Institute average MRI. The inverse solution used a local auto-regressive algorithm (LAURA) constraint (Grave de Peralta Menendez et al., 2001). The TSVD regularization (Phillips et al., 2002) was set to 10^{-3}.

The LAURA algorithm incorporates biophysical constraints regarding how rapidly an electrical field decays, such as the inverse of the square of the distance from the source. LAURA uses an autoregressive average to impose this constraint on the nearby points of each given source point. In our analysis, the radius of influence, a parameter that determines what is a nearby solution point, was set to 12.2 mm. For each solution, we focused on the regions of peak source activity.

3. Results

3.1. Behavioral results

We examined RT, accuracy, and confidence scores for lexical decisions to low-frequency, words, "rare" words, "very rare"

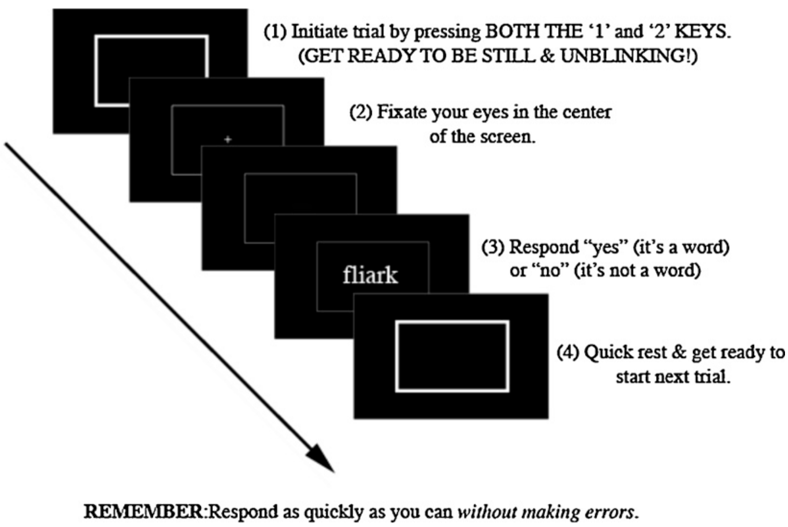


Fig. 1. Stimulus sequence.

words, and orthographically matched pseudowords. These data yielded four behavioral measures:

1. *Reaction time* (RT for correct responses in each of the four lexical categories);
2. *D-prime measures* of word–nonword discrimination (where targets are items from three real-word categories, and distracters are orthographically matched pseudowords);
3. *Self-rated confidence in meaning* (on a scale from 1 to 4) for items rated as real words (i.e., hits and false alarms); and
4. *Self-rated confidence in lexicality* for items rated as nonwords (i.e., misses and correct rejections).

A 4 (lexicality) × 3 (skill) mixed ANOVA was performed on each dependent measure. Lexicality (4 levels: low-frequency, rare, and very rare words, and pseudowords) was entered as a within-subjects variable. The group variable was skill (3 levels: high, medium, and low). Only correct responses (i.e., "hits" for real words, and "correct rejections" for nonwords) were entered into the RT analysis.

3.1.1. Response time & accuracy (d′) results

Mean RT values are shown in Fig. 2(A) below. Analyses revealed a main effect of Lexicality: $F(3, 33) = 106.11$, $p < .002$, $\varepsilon_p^2 = .94$. Post-hoc comparisons showed that responses to low-frequency words were faster than to all other categories ($p < .001$ for all comparisons). Responses to rare words tended to be faster than responses to very rare words (mean $d' = 65.53$), but this difference did not reach significance ($p = .15$). Responses to pseudowords were significantly faster than responses to very rare words ($p < .001$). The difference in response times for rare words versus pseudowords did not reach significance (mean $d' = 69.80$, $p = .08$).

There was also a main effect of skill: $F(2, 33) = 6.72$, $p = .005$, $\varepsilon_p^2 = .37$. High-skilled readers were faster than medium-skilled readers ($p = .023$) and low-skilled readers ($p = .002$). The difference between medium- and low-skilled readers was not significant ($p > .3$). There was no interaction between lexicality and skill ($p > .5$).

D-prime scores were computed separately for each participant, and for each real-word stimulus category (low-frequency, rare, and very rare). Mean *d*-prime values are shown in Fig. 2(B) below. Similar to the RT results, there was a main effect of lexicality: $F(2, 33) = 221.42$, $p < .001$, $\varepsilon_p^2 = .96$. Pairwise comparisons revealed significant differences in the average *d*-prime score for low-frequency versus rare words and for rare versus very rare ($p < .001$ for all comparisons). Subjects were most sensitive to the difference between words versus matched pseudowords for the low-frequency (familiar) items. The average *d*-prime score decreased proportionately for the rare and very rare words.

There was also a main effect of skill: $F(2, 33) = 5.19$, $p = .015$, $\varepsilon_p^2 = .34$. In contrast with the RT results, the low-skilled readers were differentiated from the high- and medium-skilled groups. There was no difference in average *d*-prime scores for medium- and high-skilled readers ($p > .2$). Hence, while average readers were just as sensitive as highly skilled readers in response to word–nonword differences, the high-skilled group was faster in making lexical decisions than both average and weaker groups. This result was further clarified by an interaction of lexicality by skill, $F(4, 66) = 4.12$, $p = .005$. While high- and average-skilled subjects showed robust discrimination of all three lexical categories, low-skilled subjects showed "floor" performance on very rare words (failing to distinguish them from pseudowords), and barely distinguished rare words from pseudowords above chance level (that is, $d' = 0$).

3.2. Stricter beta criterion for rare words in high-skilled subjects

Analysis of beta scores showed that high-skilled participants were more conservative in their lexical judgments, main effect $F(2,31) = 43.73$, $p < .05$, $\varepsilon_p^2 = .25$. More striking was the interaction between skill level and word type, interaction $F(4, 62) = 4.10$, $p < .01$, $\varepsilon_p^2 = .21$. Post-hoc comparisons showed that the main difference between skill groups was observed for rare ("frontier") words: high-skilled subjects were two to four times more conservative in their response to rare words, as compared with the average and low-skilled groups (Fig. 3).

Recall that rare words and their low-frequency counterparts do not differ along any of the orthographic or phonological dimensions investigated (see Table 2). Therefore, the differential response to rare words is likely to reflect differences in word knowledge that are semantically based. Interestingly, whatever the nature of these cues, they resulted in a different task strategy for high-skilled subjects in response to rare versus familiar and novel words.

3.2.1. Self-rated confidence

For the confidence ratings given on "word" ("yes") responses, there was a main effect of lexicality: $F(2, 33) = 126.26$, $p < .001$, $\varepsilon_p^2 = .81$. Polynomial contrasts revealed a linear trend: confidence ratings were higher for more familiar (i.e., low-frequency) words, and lower for words from the rare and very rare categories. Post-hoc comparisons further showed that mean confidence ratings for

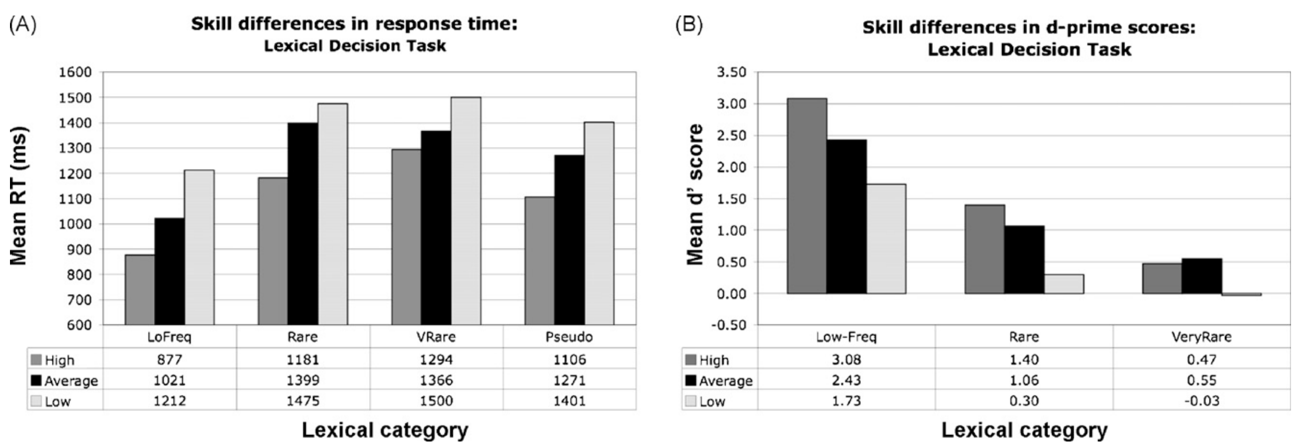


Fig. 2. (A) Mean response times and (B) mean d' scores on lexical decision task.

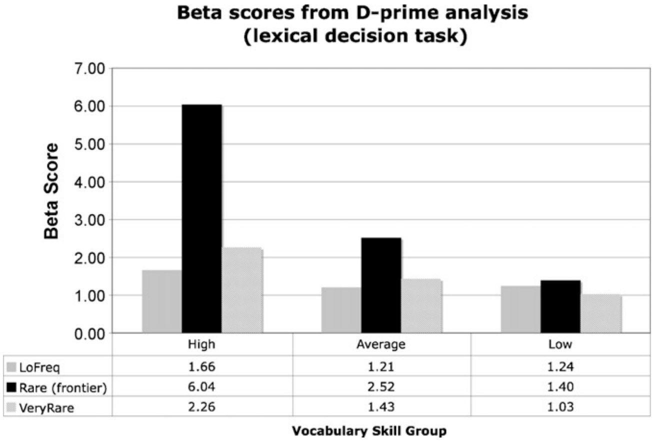


Fig. 3. Skill differences in beta criterion on the lexical discrimination task.

low-frequency words were higher than ratings for all other categories ($p < .001$), as expected. Similarly, subjects were more confident in endorsing rare versus very rare words, on average ($p < .001$). Interestingly, confidence ratings for rare words that were correctly recognized (“hits”) did not differ from confidence ratings for pseudowords that were falsely endorsed as words (“false alarms”).

Self-rated confidence for “yes” (word) responses also showed a main effect of skill, $F(2, 33) = 4.63$, $p < .05$, $\varepsilon_p^2 = .242$. Post-hoc comparisons showed that the subjects in the strong group gave higher confidence ratings than either the average or low groups, which did not differ from one another. There was no interaction of lexicality by skill on this measure.

Confidence ratings for “no” (nonword) responses showed a main effect of lexicality: $F(3, 93) = 27.84$, $p < .001$, $\varepsilon_p^2 = .79$. Paired comparisons showed that confidence ratings for pseudowords (correct rejections) were lower than for all the other categories ($p < .001$ for both comparisons). No other contrasts were significant.

Finally, there was a main effect of skill on confidence ratings for “no” trials, $F(3, 93) = 9.78$, $p < .001$. Again, high-skilled subjects gave higher confidence ratings than the other two groups.

3.3. ERP results

Grand-average ERP waveforms are shown for each of the three skill groups in Figs. 4–6. To parallel the analysis of effects of channel clusters (see Section 2), we display waveforms averaged over the electrodes in each channel group. See Appendix A for assignment of electrodes to channel groups.

As described in Section 2, the ERP data were decomposed into a set of principal (or “latent”) components. Fig. 7 summarizes the main patterns in the data, after application of PCA: ERP components are ordered from earliest to latest peak latency. The sequence of brain dynamic activity, as reflected in these scalp measures, closely replicates previous patterns observed in ERP responses to words. For example, P1 and N1 responses are initially seen over occipital cortex (over the back of the head), reflecting early visual processing, and then a sequence of left inferior cortical regions — components N1, N2, and N3 — become active as language-related processes are engaged (e.g., Frishkoff et al., 2004).

The factor scores (weights) for each principal component were entered into a mixed analysis of variance, with lexicality (4 levels: familiar words, rare hits, rare misses, and nonwords) and hemisphere (2 levels: left versus right) as within-subjects variables. Vocabulary skill (3 levels) was entered as the between-subjects factor. Factors showing significant effects of lexicality are marked with asterisks in Fig. 7. Interactions of lexicality by skill are marked by ‘+’.

3.3.1. Left posterior N2c effects

The first effects of lexicality were seen at around 250–270 ms over left posterior temporal sites. As seen in Fig. 7, this effect was captured by factor F270 (which we term the N2c, following Ritter et al., 1983). Note that the N2c has a more rostral distribution than the occipital-temporal N1 (factor F200). There was a significant effect of hemisphere, $F(1, 31) = 5.27$, $p < .05$. In addition, the three-way interaction of lexicality × hemisphere × skill approached significance, $F(1, 31) = 5.27$, $p = .055$. Post-hoc comparisons showed a tendency for high-skilled versus average and low-skilled subjects to show greater (more negative) N2c amplitudes over left posterior temporal electrodes for rare hits (i.e., rare words recognized as familiar): high versus low skilled $t(22) = -1.88$, $p = .073$, and high-skilled versus average $t(21) = -2.07$, $p = .051$. There were no group differences over the right hemisphere ($p > .2$).

Source localization of the F270 showed distributed activation along the inferior temporal pathway, with particularly strong activity in posterior ventral and medial temporal regions (Fig. 8). Consistent with the scalp topography, temporal activations are stronger in the left hemisphere.

3.3.2. Left anterior “N3” effects

An F350 factor (cf. Dien et al., 2003) showed a main effect of lexicality, $F(3, 93) = 5.52$, $p = .001$ over left frontal electrodes. As shown in Fig. 7, low-frequency words elicited less negative responses over this region. As for factor F270, activation is strongest along the left ventral stream, though the focus of activation for the N3 is more rostral (anterior IT and temporal pole).

Table 2
Stimulus attributes

	Rare (RW)	Very rare (VRW)	Known (RW match)	Known (VRW match)	Pseudowords (RW match)	Pseudowords (VRW match)
N	50	50	50	50	50	50
OLength	7.54	7.30	7.56	7.28	7.42	7.26
PLength	NR	NR	NR	NR	NR	NR
Frequency	.52	.00	3.00	3.00	–	–
Concreteness	382[a]	353[a]	344	NA[a]	–	–
Familiarity	555[a]	556[a]	411	NA[a]	–	–
Orthographic neighborhood	.30	.18	.30	.20	.30	.20
Bigram frequency	3455	3178	3440	3508	4537	4628
Phonological neighborhood	1.72	.28	1.56	2.38	.69	.46
Biphone frequency	NR	NR	NR	NR	NR	NR

[a] *Note*: not all ratings are available for each stimulus.

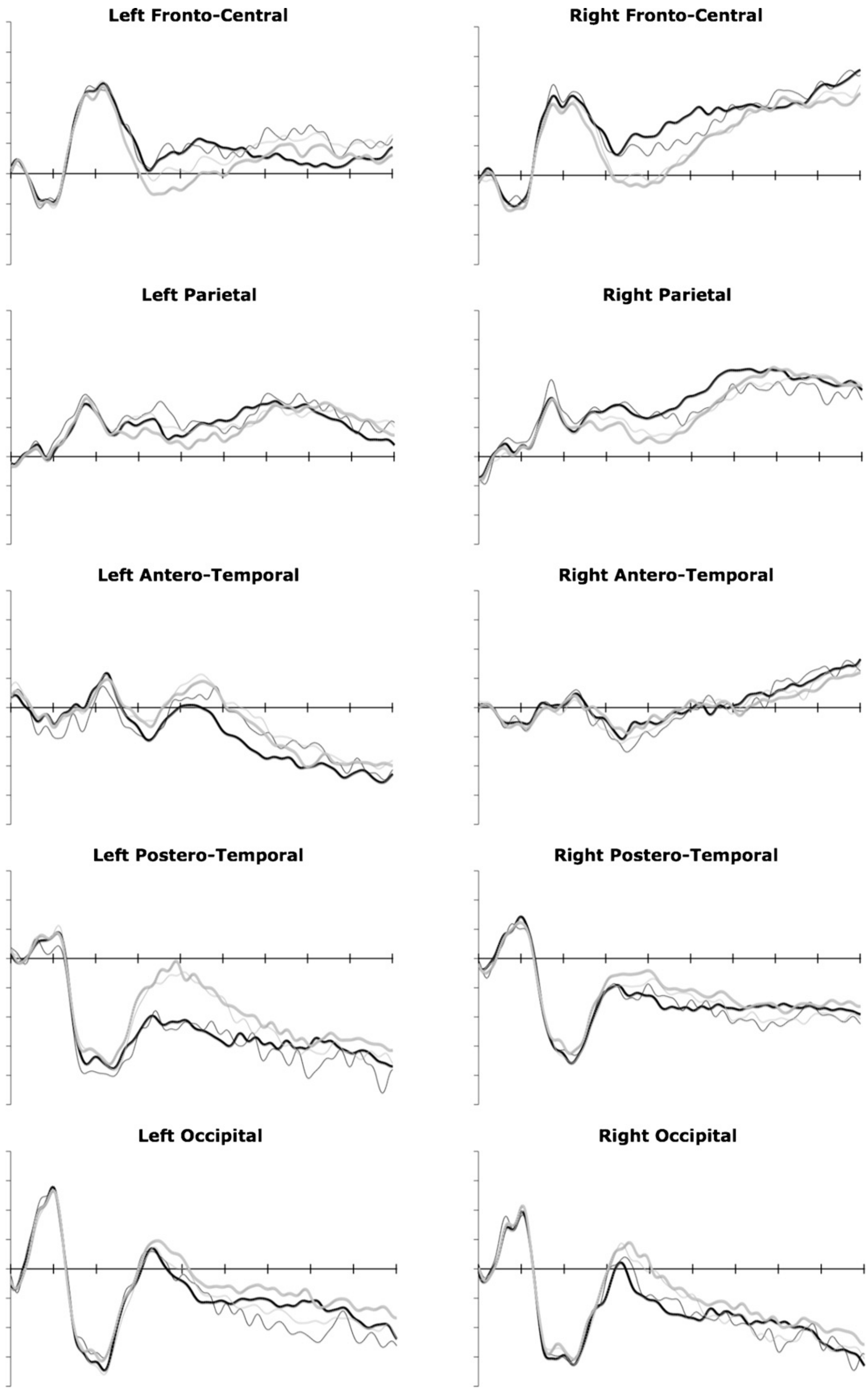


Fig. 4. Grand-averaged ERP waveforms for high-skilled subjects (n = 13). Scale, ±4 μV. Averaged over channel groups (Appendix A). Thick black line, low-frequency words (hits). Thin black line, rare words (hits); thin grey line, rare words (misses). Thick grey line, pseudowords (correct rejections).

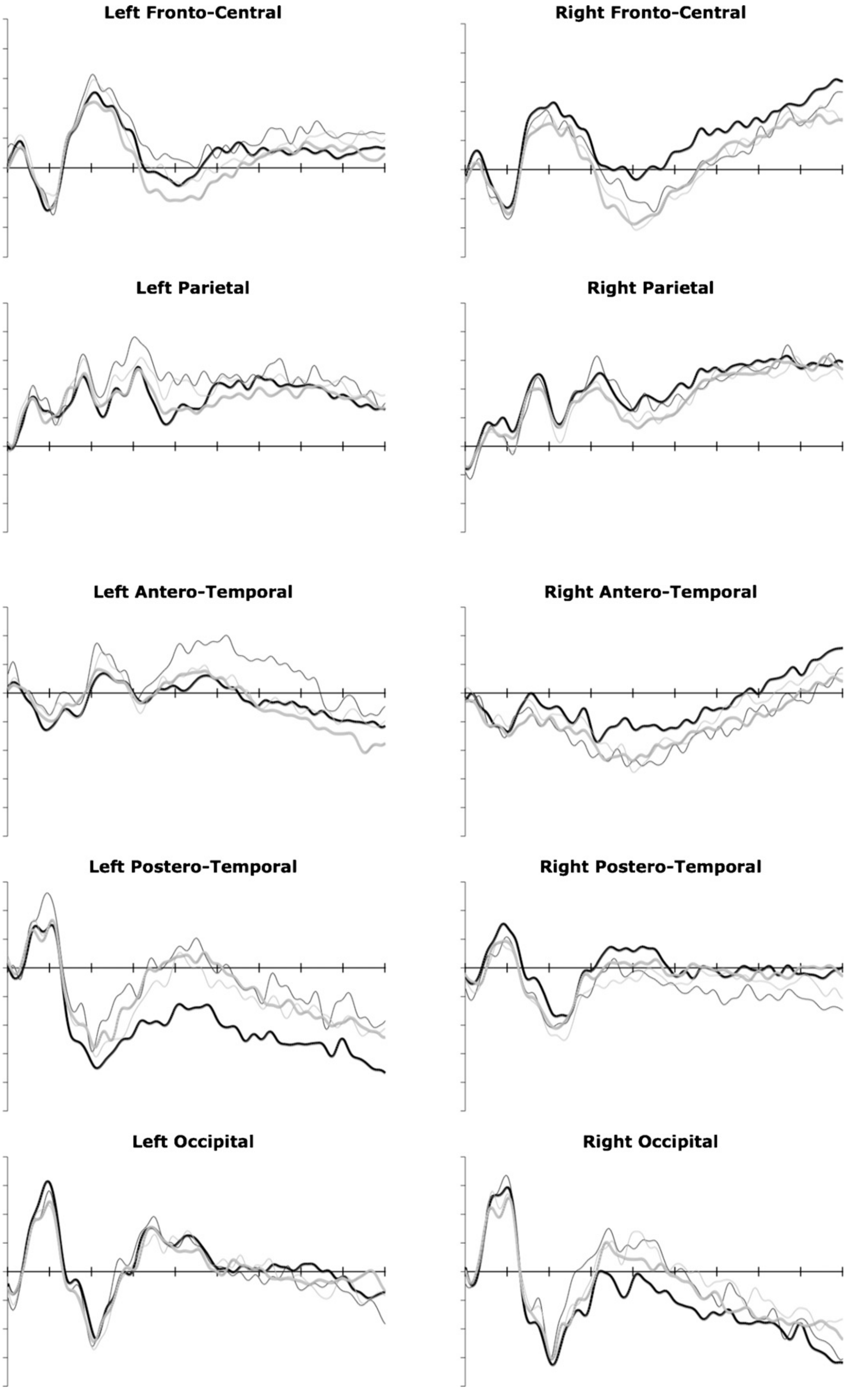


Fig. 5. Grand-averaged ERP waveforms for average-skilled subjects ($n = 10$). Scale, $\pm 4\ \mu V$. Averaged over channel groups (Appendix A). Thick black line, low-frequency words (hits). Thin black line, rare words (hits); thin grey line, rare words (misses). Thick grey line, pseudowords (correct rejections).

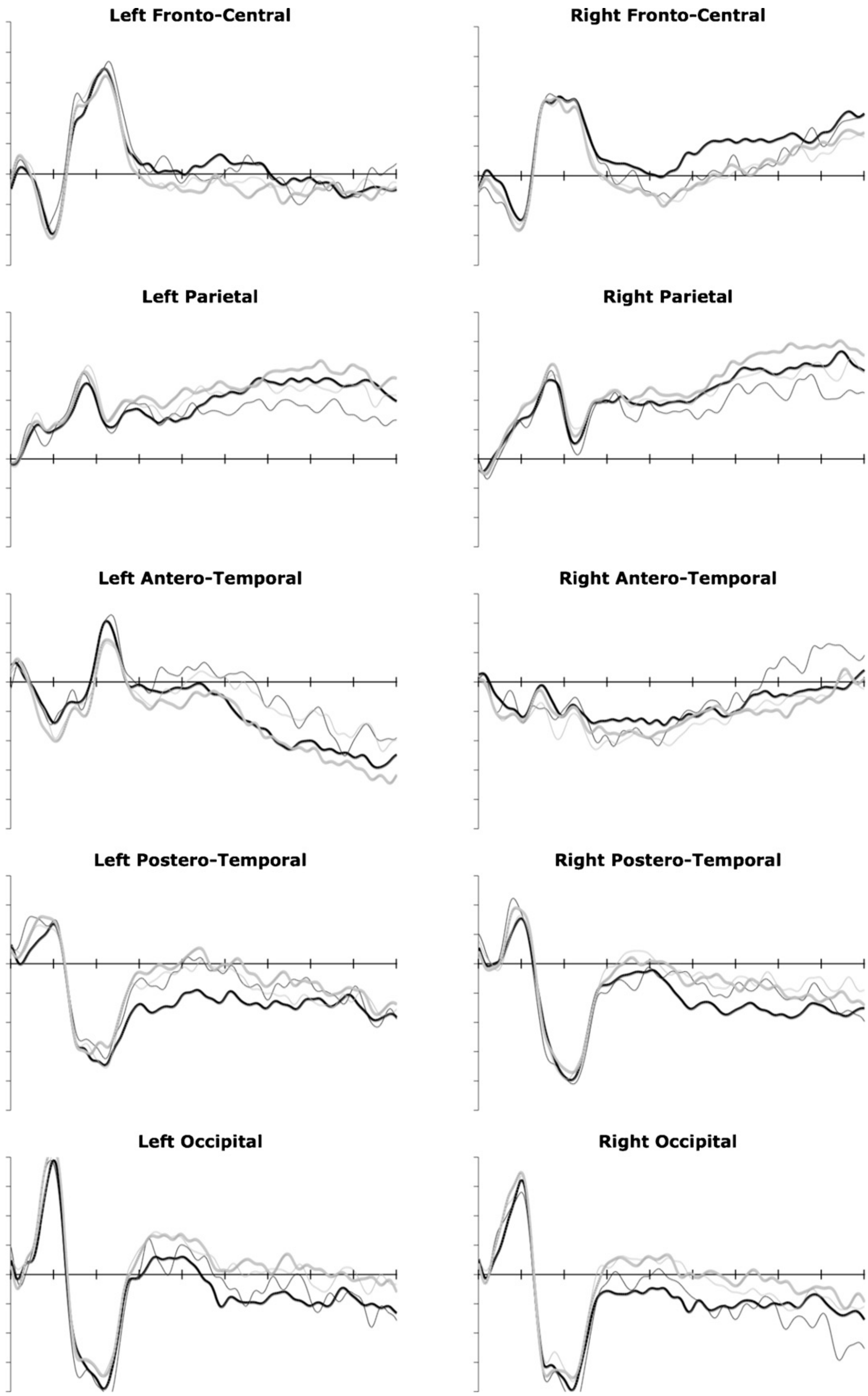


Fig. 6. Grand-averaged ERP waveforms for low-skilled subjects ($n = 11$). Scale, ±4 μV. Averaged over channel groups (Appendix A). Thick black line, low-frequency words (hits). Thin black line, rare words (hits); thin grey line, rare words (misses). Thick grey line, pseudowords (correct rejections).

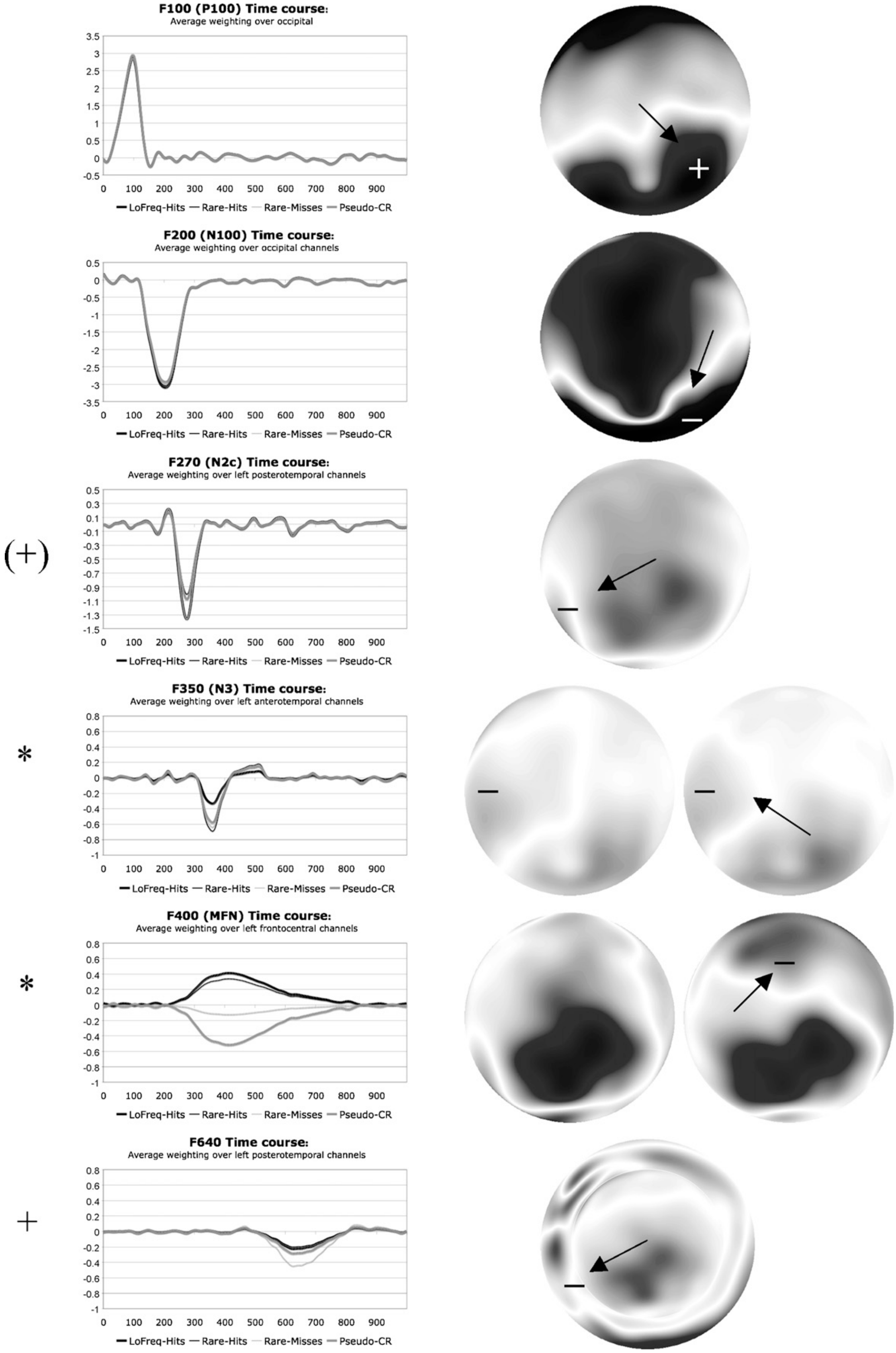


Fig. 7. Major ERP effects. Left, time course of each factor. Right, factor topographies (left, familiar words; right, pseudowords). (*) lexicality effects. (+) interactions between lexicality and skill.

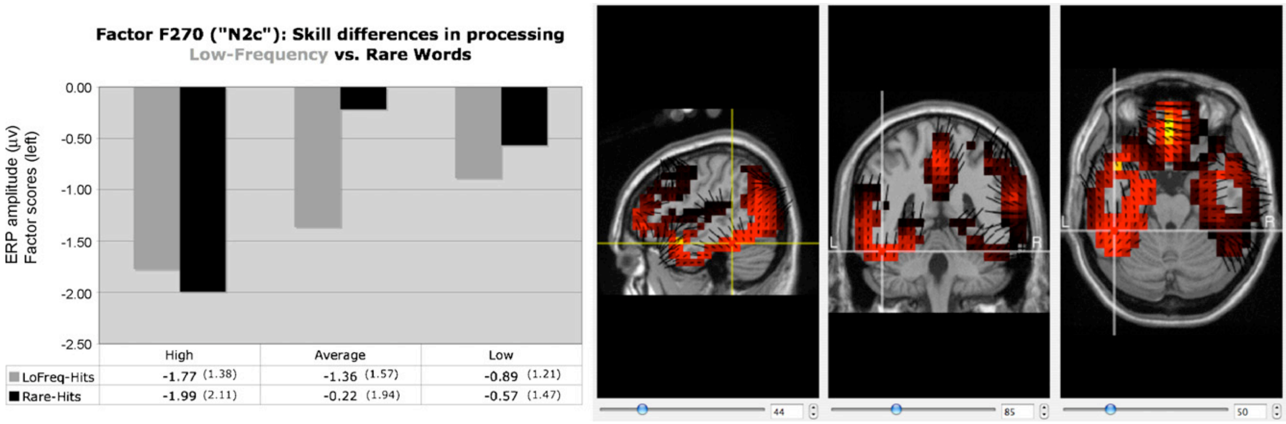


Fig. 8. Left: high-skilled, but not average or low-skilled participants, showed trend towards enhanced posterior temporal responses to rare words at around 270 ms. Right: source localization of the F270 factor points to left ventral temporal region.

3.3.3. Medial frontal "MFN" effects

A short time later, a component peaking at ~400 ms (the F400 or "MFN" factor; cf. Frishkoff, 2007, submitted for publication) shows an effect of lexicality over medial frontal areas, $F(3, 33) = 13.12$, $p < .001$. There was a strong linear trend, with less familiar stimuli evoking a greater negativity over this region.

3.3.4. Centroparietal "P3b" effects

A later component (the F550 or "P600" factor) also showed an effect of lexicality, but only over the left hemisphere, lexicality × hemisphere, $F(3, 93) = 9.04$, $p < .001$. The response to familiar words was significantly more positive over this area of the scalp, compared with all other conditions.

3.3.5. Late posterior temporal effects

Finally, factor F650, showed an interaction between lexicality and skill that was reminiscent of the pattern for the N2c (F270): $F(3,93) = 2.16$, $p = .05$. Fig. 9 shows an anterior inferior temporal focus that is stronger in the left hemisphere. As shown in Fig. 9, only high-skilled subjects showed a robust differentiation between words from the familiar (low-frequency) and rare categories over posterior temporal areas.

4. Discussion

In this experiment, we examined behavioral and ERP responses to words that vary in frequency and familiarity. Because we were particularly interested in "partial word knowledge," we selected items that occupy the lowest end of the frequency spectrum: low-frequency words, rare ("frontier") words, and very rare words, which are likely to be experienced as novel even to college-educated adults. These lexicality groups were carefully matched for orthographic (and some phonological) properties. Further, because word knowledge is likely to vary as a function of vocabulary and reading skills, we classified participants into vocabulary skill groups to examine interactions of lexicality and skill.

As predicted, participants with strong vocabulary skills showed greater fluency (faster response times) and higher accuracy on the word–nonword discrimination task. High-skilled subjects also expressed higher confidence in rating their familiarity with the meanings of words. Interestingly, analysis of beta scores indicated that the high-skilled subjects adopted a more conservative strategy, specifically in response to rare ("frontier") words, as opposed to the low-frequency and very rare words.

In addition to these behavioral outcomes, several ERP measures of brain activity showed effects of lexicality, consistent with the complexity of word semantic processing (Hill et al., 2002; Frishkoff et al., 2004; Frishkoff, 2007, submitted for publication). Most interesting was an ERP interaction effect over left temporal regions, which approached significance at 270 ms and reached significance at 640 ms: High-skilled subjects showed a specific and unique enhanced activation of left ventral temporal cortex in response to rare ("frontier") words in both time windows. These skill group differences may be related to behavioral effects strategy differences that were indicated by the beta-score analysis results.

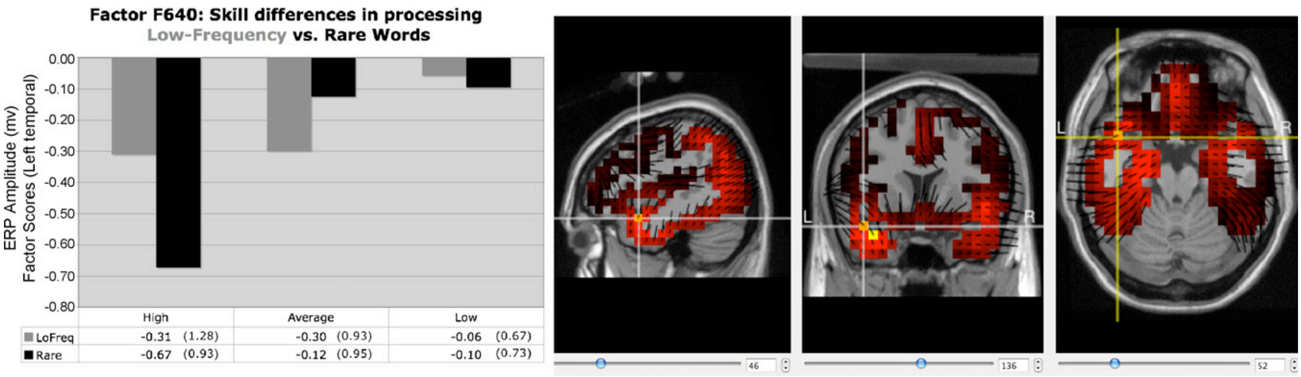


Fig. 9. Left: high-skilled, but not average or low-skilled participants, showed towards enhanced posterior temporal responses to rare words at around 640 ms. Right: source localization of the F640 factor points to left ventral temporal regions.

4.1. Early, mid-latency, and late ERP effects of lexicality

The earliest reliable effects of lexicality were seen over anterior ventral temporal regions ("N3" effect, ~350 ms). The timing and spatial distribution of this effect are similar to MEG effects described in Marinkovic et al. (2003), with increased activation for repeated versus novel words over ventral temporal areas ~250 ms in the visual modality.

While early effects of lexicality (<400 ms) are seen over the left hemisphere, later effects were seen over more medial areas of the scalp, suggesting recruitment of areas that may not be language-specific. Two ERP components, a medial frontal MFN/fN400 and a later parietal "P600" component, which have previously been linked to learning and memory processes, also showed effects of lexicality. These findings suggest an important interplay between left-hemisphere processes in word comprehension and general mechanisms of familiarity and recognition memory (cf. Curran, 2000). An important question is whether measures of familiarity, such as the MFN/fN400, reflect an automatic, unconscious response to words that have been previously seen, or whether this system is specifically sensitive to semantic memory.

At ~650 ms, there was an N2 interaction between lexicality and skill over left inferior temporal areas. High-skilled subjects, but not average or low-skilled subjects, showed a robust differentiation between familiar and less familiar words over this region. Interestingly, rare words that were recognized ("hits") were not differentiated from rare words that failed to be recognized ("misses"), even for the high-skilled subjects. Thus, although the effect is quite late (i.e., >500 ms), it appears to reflect a process that is not linked to response accuracy. It is possible that later components (perhaps closer to ~1000 ms) would differentiate rare words that were correctly recognized versus missed words. Given the mean reaction time for this task, which was much longer than in a typical lexical decision task — likely reflecting the rarity of many of the items, and the close match between words and pseudowords — it seems plausible that later components could reveal such an effect. The interaction between language-specific skills and processes and domain-general processes in learning and memory will be an important topic for future research.

4.2. Implications for individual differences in lexical quality

Word knowledge is complex, comprising multiple levels of linguistic representation, including phonological, semantic, syntactic, and — in written language — orthographic codes. On the one hand, these codes are clearly separable, leading some theories to propose distinct pathways for lexical access in reading (e.g., Coltheart et al., 2001). On the other hand, there is considerable "cross-talk" between levels of processing (Pugh et al., 2000). Indeed, the Lexical Quality Hypothesis proposes that one important difference between skilled and less-skilled readers may involve the degree to which there is efficient integration across levels of linguistic representation (Perfetti and Hart, 2001).

Previous results from Perfetti and Hart (2001, 2002) suggest that less-skilled comprehenders lack sufficient experience in reading to have developed the necessary links between form-based and semantic skills that can help to establish new form–meaning relationships. The present results add to this prior work by demonstrating that subjects with strong vocabulary skills may respond in a qualitatively different way to words that are partially known ("frontier" words): specifically, they appear less likely to endorse these items as real words, as indicated by their relatively high beta scores for rare word stimuli, relative to the other lexical categories. The explanation for this pattern is not yet clear, but it points to a greater sensitivity to items that are likely to be familiar in form, but not in meaning.

Further clues to the nature of this effect may be found in behavioral and ERP studies of word learning. Perfetti et al. (2005) recorded ERPs immediately after training participants to associate rare words (such as "gloaming") with their meaning ("twilight"). These newly learned words evoked an increased left temporal negativity at around 150–200 ms, relative to words that were either familiar or novel (rare, untrained words). The authors interpret this early negativity as an index of sublexical (orthographic) processing and note that this same effect of word training was observed in subjects with high and low reading comprehension skills. By contrast, skilled readers showed a larger increase in the P600 (P3b) effect to trained versus untrained words, as compared with less-skilled readers. This could reflect a greater "episodic memory trace," which could be related to the more successful word learning in the skilled group.

More recently, Frishkoff and Perfetti (in preparation) examined ERP markers of word learning in a task that was designed to separate effects of episodic memory (i.e., the presentation of words during training) and successful word learning. In this task, very rare words (such as "nutant") were presented in sentence contexts that either provided strong clues to the meaning of the word (high-constraint contexts), or not (low-constraint contexts). ERPs to recently trained words exhibited an increased left temporal negativity, peaking around 270–300 ms. Importantly, however, this effect was stronger for the high-constraint condition. This points to an effect that is probably not sublexical; nor is it likely to reflect episodic memory per se, since trained words occurred equally often in the high- and low-constraint conditions. Rather, this effect may reflect early lexical semantic processing, or form-to-meaning links. The timing and topography are consistent with either the left postero-temporal "N2c" or the anterior "N3" effect seen in the present study. It would be interesting to see whether skill differences similar to the ones observed in this study manifest during word learning for this left temporal component.

4.3. Implications for theories of semantic processing

Theoretically, it is of interest to know whether words that are partially known activate the same neurocognitive processes as words that are more familiar. Qualitative differences between word semantic representations could arise because stages of word learning are associated with distinct "levels," or dimensions, of meaning — such as concrete versus abstract (Paivio et al., 1988), thematic versus categorical (Ince and Christman, 2002), or global versus local semantic features (Tucker et al., 2008). Elsewhere, we have theorized that certain core semantic features (e.g., Good/Bad and Strong/Weak) may be aligned with a basic or "primitive" form of meaning representation, one that is rooted in more ancient brain mechanisms (e.g., orbitofrontal and cingulate cortex, rather than "neocortex"). The more articulated (dictionary) sense of a word may be represented by more superficial (and recently evolved) brain areas commonly implicated in language, e.g., left temporal cortex (Tucker et al., 2008). If correct, this would suggest that vague familiarity with word meaning — e.g., knowing that *invidious* has a bad connotation — may be associated with different patterns of brain activity as compared with detailed knowledge of word meanings. This would suggest that partial knowledge of word meanings represents a qualitatively different level of semantic knowledge, as compared with rich semantic representations that have formed after many encounters with a word in different contexts. Given this variability, an interesting question is whether particular dimensions of meaning are more readily acquired early in learning (e.g., concrete versus abstract, or global versus local semantic features), leading to systematic differences in the nature of representations for words that are partially known versus novel or more familiar words.

5. Conclusion

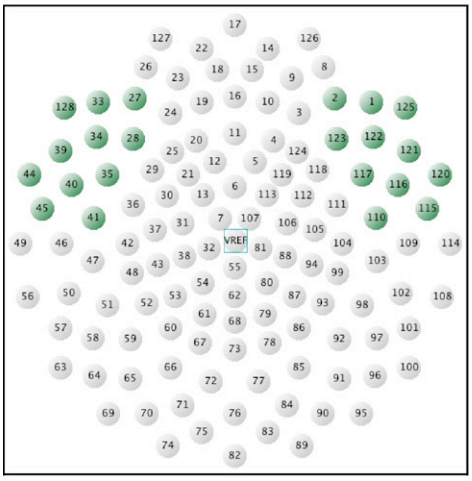
In conclusion, we describe evidence for qualitative differences in processing words that are partially known ("frontier words"), versus words that are well-known (low-frequency) or completely novel (extremely rare). These differences were unique to subjects with strong vocabulary skills. A d'/beta analysis of word–nonword discrimination was consistent with a different strategy in processing frontier words in the high-skilled group. Moreover, high-skilled, but not average or low-skilled, participants showed a trend towards enhanced inferior temporal "N2" effects, and the same pattern of effects was observed over left inferior temporal sites again at ~650 ms. Source localization pointed to generators in left inferior temporal regions. We speculate that the posterior temporal effects reflect enhanced attention to word form, which is elicited by top-down processing, due to feedback from frontal to temporal regions. This proposal will need to be tested in future studies.

Acknowledgements

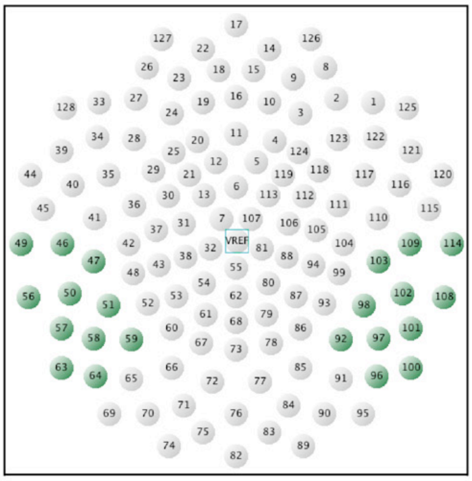
This research was supported by an American Psychological Association/Institute of Education Sciences Postdoctoral Education Research Training fellowship under the Department of Education, grant number R305U030004. We thank Jenifer Einstein, Jaimee Shepard, and Natalie Campbell for assistance with data acquisition.

Appendix A. Electrode clusters

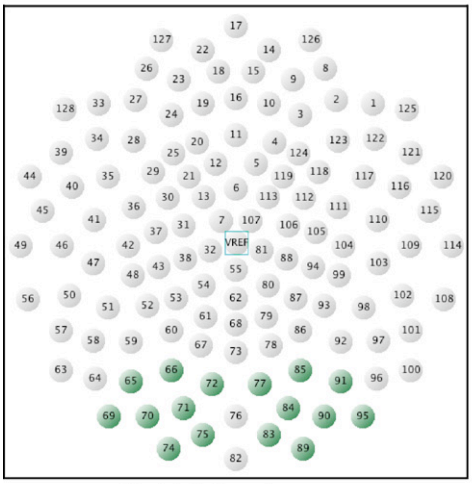
Appendix A: Electrode Clusters



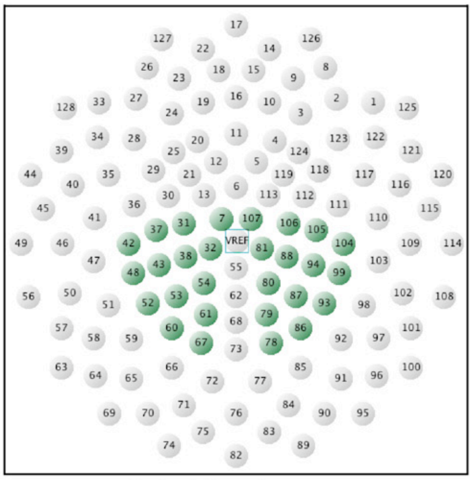
Anterior temporal channels



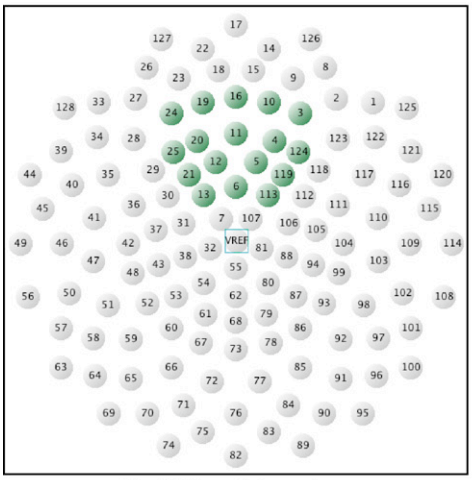
Posterior temporal channels



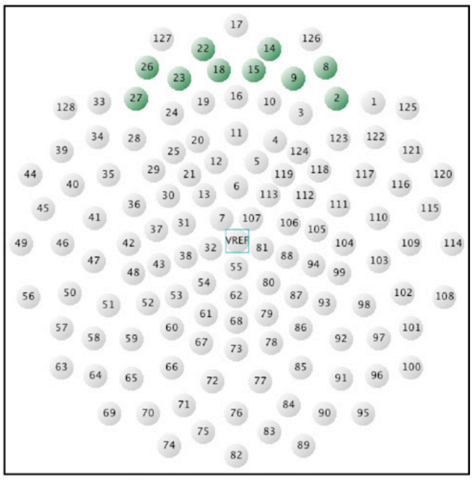
Occipital channels



Parietal channels



Medial frontal channels



Orbitofrontal channels

Appendix B. Stimuli

MatchNo	StimNo	Familiar word	StimNo	Very rare	Synonym	StimNo	Nonword
001	001	Abject	151	Nutant	Drooping	201	Aboats
002	002	Cryptic	152	Aphotic	Dark	202	Ingraid
003	003	Forlorn	153	Operose	Difficult	203	Angside
004	004	Colossal	154	Maenadic	Crazy	204	Aprobble
005	005	Amenable	155	Latitent	Hidden	205	Punmater
006	006	Staunch	156	Pertuse	Torn	206	Breerse
007	007	Septic	157	Arrant	Obvious	207	Bummit
008	008	Guttural	158	Plangent	Ringing	208	Frakle
009	009	Venereal	159	Caritive	Lacking	209	Veebrate
010	010	Astray	160	Arrect	Upright	210	Deeple
011	011	Defunct	161	Pollent	Strong	211	Desheen
012	012	Expectant	162	Cautelous	Careful	212	Pandlings
013	013	Mimetic	163	Torvous	Stern	213	Kooshle
014	014	Sublime	164	Ringent	Gaping	214	Quarood
015	015	Ebullient	165	Aleatoric	Chance	215	Grevelly
016	016	Psychic	166	Priscan	Ancient	216	Athress
017	017	Nocturnal	167	Thrasonic	Boastful	217	Parkenize
018	018	Ponderous	168	Precative	Yielding	218	Rankorens
019	019	Nebulous	169	Mofussil	Rural	219	Ordition
020	020	Ethereal	170	Oragious	Stormy	220	Greewood
021	021	Vascular	171	Agminate	Collected	221	Hinsnash
022	022	Erudite	172	Vafrous	Sly	222	Imsteed
023	023	Wanton	173	Whelky	Round	223	Infees
024	024	Frontal	174	Zetetic	Seeking	224	Inprary
025	025	Idiotic	175	Vagient	Crying	225	Noskings
026	026	Cortical	176	Orgulous	Proud	226	Shruping
027	027	Migratory	177	Accolent	Near	227	Inlumant
028	028	Sordid	178	Dumose	Bushy	228	Unstab
029	029	Flagrant	179	Discinct	Undressed	229	Nitarize
030	030	Mundane	180	Decumen	Main	230	Spafoke
031	031	Velvety	181	Ennomic	Legal	231	Oliteed
032	032	Pallid	182	Curule	High	232	Onally
033	033	Fugitive	183	Draffish	Worthless	233	Palition
034	034	Tolerable	184	Conticent	Silent	234	Wenthing
035	035	Prosodic	185	Eldritch	Occult	235	Portlany
036	036	Arable	186	Xeric	Dry	236	Peeda
037	037	Congruent	187	Infandous	Evil	237	Chekshing
038	038	Erratic	188	Impavid	Brave	238	Solants
039	039	Brittle	189	Iracund	Upset	239	Spamine
040	040	Odious	190	Hulchy	Swollen	240	Taksub
041	041	Stubby	191	Colytic	Preventive	241	Straret
042	042	Illicit	192	Loxotic	Distorted	242	Sustabs
043	043	Demure	193	Incult	Rude	243	Laitle
044	044	Rustic	194	Irenic	Peaceful	244	Tarler
045	045	Reverent	195	Irrisory	Joking	245	Teebored
046	046	Molten	196	Hiemal	Wintry	246	Rafide
047	047	Transient	197	Fingent	Shapeless	247	Tobeese
048	048	Proximate	198	Esculent	Edible	248	Breezeese
049	049	Dynastic	199	Epinosic	Toxic	249	Mortions
050	050	Astride	200	Mugient	Bellowing	250	Randans
051	051	Mystic	101	Lenitive	Soothing	251	Masaitor
052	052	Ecliptic	102	Addled	Confused	252	Lounass
053	053	Imminent	103	Effete	Weak	253	Partant
054	054	Deductive	104	Meracious	Scary	254	Fewnate
055	055	Devilish	105	Macilent	Thin	255	Inverst
056	056	Lucrative	106	Recondite	Obscure	256	Arbling
057	057	Watery	107	Limpid	Clear	257	Arkant
058	058	Kindred	108	Raffish	Wild	258	Asteest
059	059	Doctrinal	109	Vitriolic	Hateful	259	Foanst
060	060	Cellular	110	Unctious	Oily	260	Bideroun
061	061	Epileptic	111	Refulgent	Shiny	261	Bumparize
062	062	Abortive	112	Abducent	Turning	262	Chertary
063	063	Systemic	113	Fungible	Same	263	Kormants
064	064	Versatile	114	Saturnine	Sad	264	Faitioner
065	065	Soluble	115	Condign	Suitable	265	Franned
066	066	Payable	116	Risible	Humorous	266	Freekan
067	067	Freudian	117	Salacious	Lustful	267	Grections
068	068	Flashy	118	Turgid	Showy	268	Grooth
069	069	Pitiable	119	Abditive	Secret	269	Hampling
070	070	Sunken	120	Otiose	Useless	270	Hapsty
071	071	Ablaze	121	Tetchy	Angry	271	Imtate
072	072	Satirical	122	Truculent	Cruel	272	Indanlize
073	073	Migrant	123	Lissome	Pliable	273	Groudal
074	074	Lenient	124	Pinguid	Fat	274	Crosile
075	075	Hesitant	125	Inimical	Opposed	275	Laindent

Appendix B (*Continued*)

MatchNo	StimNo	Familiar word	StimNo	Very rare	Synonym	StimNo	Nonword
076	076	Crummy	126	Torpid	Inactive	276	Oaline
077	077	Granular	127	Fastuous	Rich	277	Dultions
078	078	Insecure	128	Minatory	Sinister	278	Pefitant
079	079	Candid	129	Prolix	Wordy	279	Prakid
080	080	Wayward	130	Bilious	Irritable	280	Pralics
081	081	Eclectic	131	Ancillary	Related	281	Prastramp
082	082	Adorable	132	Specious	False	282	Promming
083	083	Titanic	133	Lambent	Glowing	283	Proolit
084	084	Hectic	134	Jocose	Funny	284	Proosh
085	085	Impudent	135	Bibulous	Drunk	285	Puntlize
086	086	Defiant	136	Maudlin	Tearful	286	Raitral
087	087	Reddish	137	Reboant	Loud	287	Rethads
088	088	Ludicrous	138	Impacable	Hostile	288	Samanning
089	089	Dainty	139	Nocent	Harmful	289	Satate
090	090	Conjugal	140	Incondite	Flawed	290	Scruside
091	091	Fragrant	141	Immanent	Essential	291	Sorabble
092	092	Motley	142	Vorant	Hungry	292	Spizma
093	093	Slovenly	143	Apposite	Fitting	293	Sprabble
094	094	Vivacious	144	Fractious	Unruly	294	Spratling
095	095	Unwise	145	Labile	Flexible	295	Sylner
096	096	Platonic	146	Inchoate	Unfinished	296	Taitabel
097	097	Deficient	147	Invidious	Offensive	297	Traitrets
098	098	Observant	148	Verecund	Shy	298	Untannist
099	099	Sceptical	149	Trenchant	Forceful	299	Varndifle
100	100	Neural	150	Lepid	Pleasant	300	Raran

References

Beck, I., McKeown, M., Kucan, L., 2002. Bringing Words to Life: Robust Vocabulary Instruction. Guilford Press, New York.

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 (3), 235–260.

Binder, J.R., McKiernan, K.A., Parsons, M.E., Westbury, C.F., Possing, E.T., Kaufman, J.N., et al., 2003. Neural correlates of lexical access during visual word recognition. Journal of Cognitive Neuroscience 15 (3), 372–393.

Brown, J.I., Fishko, V.V., Hanna, G., 1993. Nelson–Denny Reading Test. Riverside Publishing Co., New York.

Brown, J., Frishkoff, G., Eskenazi, M., 2005. Automatic question generation for vocabulary assessment. In: Proceedings of the Annual Human Language Technology Meeting, October 6–8, Vancouver, BC.

Cain, K., Oakhill, J.V., Elbro, C., 2003. The ability to learn new word meanings from context by school-age children with and without language comprehension difficulties. Journal of Child Language 30 (3), 681–694.

Chrisomalis, S.A., 2005. The phrontistry: English word lists and language resources. http://www.kamous.com/translator/s.asp?l=1656. Accessed on November 1, 2005.

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.

Curran, T., 2000. Brain potentials of recollection and familiarity. Memory & Cognition 28 (6), 923–938.

Dale, E., 1965. Vocabulary measurement: techniques and major findings. Elementary English 42, 82–88.

Dien, J., 2004. *PCA toolbox* (Version 1.093). Lawrence, KA.

Dien, J., Beal, D.J., Berg, P., 2005. Optimizing principal components analysis of event-related potentials: matrix type, factor loading weighting, extraction, and rotations. Clinical Neurophysiology 116 (8), 1808–1825.

Dien, J., Frishkoff, G.A., 2005. Introduction to principal components analysis of event-related potentials. In: Handy, T. (Ed.), Event-Related Potentials: A Methods Handbook. MIT Press, Cambridge, MA, pp. 189–208.

Dien, J., Frishkoff, G.A., Cerbone, A., Tucker, D.M., 2003. Parametric analysis of event-related potentials in semantic comprehension: evidence for parallel brain mechanisms. Brain Research Cognitive brain Research 15 (2), 137–153.

Durso, F.T., Shore, W.J., 1991. Partial knowledge of word meanings. Journal of Experimental Psychology: General 120, 190–202.

Frishkoff, G.A., 2007. Hemispheric differences in strong versus weak semantic priming: evidence from event-related brain potentials. Brain & Language 100 (1), 23–43.

Frishkoff, G.A., submitted for publication. Fractionating the N400 Effect: separation of frontal and posterolateral responses in semantic priming. Clinical Neurophysiology.

Frishkoff, Collins-Thompson, Perfetti, Callan, in press. Measuring incremental changes in word knowledge: Experimental validation and implications for learning and assessment. Behavioral Research Methods.

Frishkoff, Perfetti, in preparation. Incremental learning of word meaning from context: effects on MFN, N3 & N400 components.

Frishkoff, G.A., Tucker, D.M., Davey, C., Scherg, M., 2004. Frontal and posterior sources of event-related potentials in semantic comprehension. Cognitive Brain Research 20 (3), 329–354.

Grave de Peralta Menendez, R., Gonzalez Andino, S., Lantz, C., Michel, C., Landis, T., 2001. Noninvasive localization of electromagnetic epileptic activity. I. Method descriptions and simulations. Brain Topography 14, 131–137.

Hill, H., Strube, M., Roesch-Ely, D., Weisbrod, M., 2002. Automatic vs. Controlled processes in semantic priming — differentiation by event-related potentials. International Journal of Psychophysiology 44 (3), 197–218.

Hinojosa, J.A., Martin-Loeches, M., Munoz, F., Casado, P., Fernandez-Frias, C., Pozo, M.A., 2001. Electrophysiological evidence of a semantic system commonly accessed by animals and tools categories. Cognitive Brain Research 12 (2), 321–328.

Ince, E., Christman, S.D., 2002. Semantic representations of word meanings by the cerebral hemispheres. Brain & Language 80 (3), 393–420.

Kounios, J., Holcomb, P.J., 1994. Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. Journal of Experimental Psychology: Learning Memory and Cognition 20 (4), 804–823.

Kutas, M., Federmeier, K.D., 2000. Electrophysiology reveals semantic memory use in language comprehension. Trends in Cognitive Science 4 (12), 463–470.

Marinkovic, K., Dhond, R.P., Dale, A.M., Glessner, M., Carr, V., Halgren, E., 2003. Spatiotemporal dynamics of modality-specific and supramodal word processing. Neuron 38 (3), 487–497.

Martin-Loeches, M., Hinojosa, J.A., Casado, P., Munoz, F., Fernandez-Frias, C., 2004. Electrophysiological evidence of an early effect of sentence context in reading. Biological Psychology 65 (3), 265–280.

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

Nagy, W.E., Anderson, R.C., Herman, P.A., 1987. Learning words from context during normal reading. American Educational Research Journal 24 (2), 237–270.

Neville, H.J., Mills, D.L., Lawson, D.S., 1992. Fractionating language: different neural subsystems with different sensitive periods. Cerebral Cortex 2 (3), 244–258.

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

Paivio, A., Clark, J.M., Khan, M., 1988. Effects of concreteness and semantic relatedness on composite imagery ratings and cued recall. Memory & Cognition 16 (5), 422–430.

Perfetti, C.A., Hart, L., 2001. The lexical basis of comprehension skill. In: Gorfien, D.S. (Ed.), On the Consequences of Meaning Selection: Perspectives On Resolving Lexical Ambiguity. American Psychological Association, Washington, DC, pp. 67–86.

Perfetti, C.A., Hart, L., 2002. The lexical quality hypothesis. In: Verhoeven, L., Elbro, C., Reitsma, P. (Eds.), Precursors of Functional Literacy, vol. 11. John Benjamins, Amsterdam, pp. 67–86.

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, & Cognition 31 (6), 1281–1292.

Pexman, P.M., Hargreaves, I.S., Edwards, J.D., Henry, L.C., Goodyear, B.G., 2007. Neural correlates of concreteness in semantic categorization. Journal of Cognitive Neuroscience 19 (8), 1407–1419.

Phillips, C., Rugg, M.D., Fristont, K.J., 2002. Systematic regularization of linear inverse solutions of the EEG source localization problem. Neuroimage 17 (1), 287–301.

Proverbio, A.M., Adorni, R., Zani, A., in press. Inferring native language from early bio-electrical activity.

Pugh, K.R., Mencl, W.E., Shaywitz, B.A., Shaywitz, S.E., Fulbright, R.K., Constable, R.T., et al., 2000. The angular gyrus in developmental dyslexia: task-specific differences in functional connectivity within posterior cortex. Psychological Science 11 (1), 51–56.

Raven, J.C., 1960. Guide to the Standard Progressive Matrices. H.K. Lewis, London.

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 (3), 219–237.

Ritter, W., Simson, R., Vaughan, H.G., 1983. Event-related potential correlates of two stages of information processing in physical and semantic discrimination tasks. Psychophysiology 20 (2), 168–179.

Rossell, S.L., Price, C.J., Nobre, A.C., 2003. The anatomy and time course of semantic priming investigated by fMRI and ERPs. Neuropsychologia 41 (5), 550–564.

Rudell, A.P., 1992. Rapid stream stimulation and the recognition potential. Electroencephalography & Clinical Neurophysiology 83 (1), 77–82.

Rudell, A.P., 1999. The recognition potential and the word frequency effect at a high rate of word presentation. Cognitive Brain Research 8 (2), 173–175.

Rugg, M.D., Curran, T., 2007. Event-related potentials and recognition memory. Trends in Cognitive Science 11 (6), 251–257.

Silva-Pereyra, J., Rivera-Gaxiola, M., Aubert, E., Bosch, J., Galan, L., Salazar, A., 2003. N400 during lexical decision tasks: a current source localization study. Clinical Neurophysiology 114 (12), 2469–2486.

Stahl, S.A., 1986. Three principles of effective vocabulary instruction. Journal of Reading 29 (7), 662–668.

Stanovich, K.E., West, R.F., 1989. Exposure to print and orthographic processing. Reading Research Quarterly 24, 402–433.

Swanborn, M.S.L., De Glopper, K., 2002. Impact of reading purpose on incidental word learning from context. Language Learning 52 (1), 95–117.

Tucker, D.M., 2002. Embodied meaning: an evolutionary–developmental analysis of adaptive semantics. In: Givon, T., Malle, B. (Eds.), The Evolution of Language out of Pre-Language. 53rd ed. J. Benjamins, Amsterdam, pp. 51–82.

Tucker, D.M., Frishkoff, G.A., Luu, P., 2008. Microgenesis of language. In: Stemmer, B., Whitaker, H.A. (Eds.), Handbook of Neuroscience of Language. Elsevier, Cambridge, MA, pp. 45–56.

Tucker, D.M., Luu, P., Frishkoff, G., Quiring, J., Poulsen, C., 2003. Frontolimbic response to negative feedback in clinical depression. Journal of Abnormal Psychology 112 (4), 667–678.

Van Petten, C., Luka, B.J., 2006. Neural localization of semantic context effects in electromagnetic and hemodynamic studies. Brain and Language 97 (3), 279–293.

Westbury, C., Buchanan, L., 2002. The probability of the least likely non-length-controlled bigram affects lexical decision reaction times. Brain & Language 81 (1–3), 66–78.