Available online at www.sciencedirect.com

ScienceDirect

Brain and Language 102 (2007) 30–45

Brain and Language

www.elsevier.com/locate/b&l

# An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders

Nicole Landi *, Charles A. Perfetti

*University of Pittsburgh, Learning Research and Development Center and Center for the Neural Basis of Cognition, Pittsburgh, PA, USA*

Accepted 1 November 2006
Available online 22 December 2006

## Abstract

The most prominent theories of reading consider reading comprehension ability to be a direct consequence of lower-level reading skills. Recently however, research has shown that some children with poor comprehension ability perform normally on tests of lower-level skills (e.g., decoding). One promising line of behavioral research has found semantic processing differences between good and poor comprehenders and suggests that impoverished semantic ability may be linked to poor comprehenders' difficulties. In the current study, we used event related potentials (ERP) to compare adult skilled and less-skilled comprehenders on a set of semantic and phonological processing tasks. The results revealed that the N400 component of the ERP and the P200 component were sensitive to differences between skilled and less-skilled comprehenders during a semantic processing task. Importantly, skilled and less-skilled comprehenders showed no differences in their ERP response during a phonological processing task. These findings provide neurophysiological support for the hypothesis that less-skilled comprehenders have a weakness in semantic processing that may contribute to their comprehension difficulties.

*Keywords:* Reading comprehension; Semantics; Phonology; Reading skill; ERP

## 1. Introduction

### 1.1. Semantic processing in poor comprehenders

It is now well established that less-skilled comprehenders are more likely than their peers to have difficulty with conversion of graphemic information into phonemic information (Beck & Juel, 1992; Vellutino & Scanlon, 1987). This finding has led to causal hypotheses that link low level lexical processing skills (particularly decoding) to comprehension skill (e.g. verbal efficiency theory, Perfetti, 1985). One question left open by these early hypotheses was the nature of the lexical processing difficulties that limit comprehension, especially whether meaning-level lexical processes limit comprehension beyond decoding and word identification. For example Perfetti (1985) found that children with poor comprehension skill showed slow meaning retrieval in a semantic search task, suggesting difficulty beyond slow word decoding. Furthermore, Perfetti, Hogaboam, and Bell, cited in Perfetti and Lesgold (1979) found that children who were poor comprehenders were slower in a semantic categorization task but not in a simple word matching task. These studies suggested semantic processing difficulties beyond word reading. However, in these studies, children with comprehension difficulties also had decoding difficulties, making conclusions about the relationship between semantic processing, decoding and comprehension difficult.

Recent studies have identified populations of readers with comprehension difficulty in the absence of decoding problems (Cain, Oakhill, & Bryant, 2000; Oakhill & Cain, 2000; Oakhill, Cain, & Bryant, 2003). On some estimates, 10% of children in the 7–10 year old age range have this specific problem, which has been termed "specific comprehension impairment" (Nation, 2005; Stothard & Hulme, 1995). Such children have been reported to have trouble

---

* Corresponding author. Present address: Haskins Laboratories, 300 George Street, New Haven, CT 06511, USA.
*E-mail address:* Landi@haskins.yale.edu (N. Landi).

0093-934X/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.bandl.2006.11.001

generating text-appropriate inferences (Oakhill & Cain, 2000), monitoring their comprehension progress (Oakhill et al., 2003), using relevant semantic information (Nation & Snowling, 1998), processing syntactically complex sentences (Hagtvet, 2003), and holding information online in working memory (Oakhill et al., 2003); all despite seemingly normal decoding ability.

One hypothesis about the nature of this impairment is that these poor comprehenders may have general semantic processing difficulties (Nation & Snowling, 1998, 1999). Support for this hypothesis comes from evidence that poor comprehenders perform more poorly than controls on a variety of semantic tasks. For example, Nation and Snowling (1998) found that readers with comprehension impairments were significantly slower than controls at reading exception and low frequency words—words that are thought to require greater input from semantics, but not at reading regular high frequency words, suggesting normal decoding skills. Poor comprehenders also generated fewer semantic category members in a verbal fluency task relative to age match controls (Nation & Snowling, 1998) and had trouble judging synonyms relative to controls, demonstrating difficulty with receptive as well as productive semantics (Nation & Snowling, 1998, 1999). Furthermore, poor comprehenders showed selective difficulty for low imageability or abstract words relative to concrete words in semantic judgment and recall tasks (Nation & Snowling, 1999) and they had poorer-than-normal performance naming low-frequency pictures (Nation, Marshall, & Snowling, 2001). These findings mirror some of the difficulties found in individuals with semantic impairments[1] (Barr & Brandt, 1996; Bird, Lambon Ralph, Patterson, & Hodges, 2000; Patterson & Behrmann, 1997; Patterson & Hodges, 1992; Ward, Stott, & Parkin, 2000).

Research has also suggested that this semantic processing difficulty in poor comprehenders may be specific to certain types of semantic relationships. For example, Nation and Snowling (1999) found that poor comprehenders showed semantic priming for categorically related items that are highly associated (e.g., cat–dog) and/or have functional relationships (e.g., hammer–nail) but no priming for categorically related items that are not commonly associated or functionally related (e.g., nose–head). Control adults and children showed priming for all pair types, suggesting that the connections between categorically related items may be weaker for less-skilled comprehenders. Nation and Snowling suggested that associative relationships may be bolstered by lexical and real world co-occurrence and are thus more robust whereas categorical relationships emerge out of increasingly refined experiences and are therefore relatively late developing (but see Cree & McRae, 2003, for a different account of the basis of these relationships).

One additional question is whether these semantic difficulties may represent a developmental delay or lag that would disappear with age or a more basic deficit that would resist a developmental catch up. The developmental delay view is consistent with the observation that abstract semantic relationships tend to develop late (Nation & Snowling, 1999). The delay account is also consistent with the finding that surface dyslexic children's "semantic" word reading errors (poor exception word reading) are minimized when compared to reading age controls rather than chronological age controls, but phonological dyslexic children's phonological errors (poor non-word reading) are found in both comparisons (Manis, Seidenberg, Doi, McBride-Chang, & Peterson, 1996; Stanovich, Siegel, & Gottardo, 1997). On the other hand, Nation (2005) reported that 78% of poor comprehenders tested at age 8 and 9 still had comprehension impairments when tested at age 13 and 14, suggesting that the impairment is persistent.

The persistence of comprehension-specific problems into adulthood is supported by a study by Landi (2006), who found dissociations between comprehension skill and low-level reading skills also exist in adult readers. Using principle components analysis (PCA), Landi demonstrated that low-level reading (decoding) was dissociable from comprehension and other higher-level reading skills such as vocabulary knowledge in adults. That is, the two sets of skills clustered into two non-overlapping components. Furthermore, by comparing the number of individuals who scored high on the high-level reading skill component vs. the low-level reading skill component, she found that although the majority of the participants scored well on both or on neither of the components, a number of participants had a discrepancy between their two scores. In particular, 187 out of 799 people (23%) scored above average on the low-level component and below average for the high-level component; only 68 (9%) of the participants had the reverse discrepancy. These findings suggest that the general dissociation between high-level and low-level reading abilities occurs in adults, perhaps even in larger numbers than in children. This increase in specific high-level impairment follows naturally from the greater difficulty of comprehension relative to decoding. Decoding is governed by a finite set of phoneme–grapheme regularities, but semantics is an unlimited and largely arbitrary domain.

Further testing of semantic performance with adults would help clarify whether the comprehension problems of adults may also be caused by a semantic processing difficulty, suggesting a general link between comprehension skill and semantic skill.

### *1.2. Electrophysiological markers of semantic and phonological processing*

One way we can help elucidate potential underlying processing differences between groups is to use a direct

---

[1] It should be noted that unlike poor comprehenders, individuals with semantic impairments have known cortical damage, thus any parallels in behavioral performance must be interpreted with some caution (also see Thomas & Karmaloff-Smith, 2002, for a cautionary take on comparing developmental differences to brain damage).

measure of neural processing such as ERP. ERP's are useful for revealing underlying processing differences in a variety of tasks because they provide a continuous record of brain activity from the beginning of stimulus onset that is temporally accurate to the millisecond and thus not confounded with post processing differences (e.g. decision-making). For example, researchers have found early components (100–200 ms) that correspond to processing of early orthographic components of words, somewhat later components (200–500 ms) that correspond to phonological and semantic processing of words and late (600 ms) components that correspond to sentence or syntactic level processing (Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier, 1999; Kramer & Donchin, 1987; Kutas & Van Petten, 1994; Rugg, 1984; Perez-Abalo, Rodriguez, Bobes, Gutierrez, & Valdes-Sosa, 1994). Thus, comparing skilled and less-skilled comprehenders ERP's in phonological and semantic tasks may help inform our understanding of specific comprehension impairments by providing estimates of processing abilities in the two groups that are temporally and neurophysiologically precise.

The N400 component is of particular relevance for the current study because it is a component that is sensitive to both semantic and phonological processing differences. The N400 is a scalp negative waveform (or relative negativity) that is typically largest in central and parietal regions and is associated with relatedness. For example, in categorization tasks or semantic priming tasks, the N400 is larger (more negative) to a target stimulus that is unrelated or less-related to a probe stimulus compared with trials where the target word and probe word are related. This effect is also observed for semantically inconsistent words that appear in sentences (e.g., "The food was too hot to eat/swim"; in this case the N400 would be larger when the participant read the sentence with the word "swim" than when they read the sentence with the word "eat"; see Hillyard & Kutas, 2002, for a review).

Although the N400 response to semantic relatedness is extremely robust, this component is sensitive to phonological relations as well. Several studies have shown that the N400 is larger for phonologically distinct relative to phonologically similar stimulus pairs a rhyme task (larger N400 for non-rhyming stimuli) and in a phonological oddball detection task (Kramer & Donchin, 1987; Radeau, Besson, Fonteneau, & Castro, 1998). These findings suggest that the N400 discriminates between words in a pair that are similar along some dimension compared to those that are dissimilar.

Although the N400 is the most consistently identified component in phonological tasks, several studies have identified earlier components that are also sensitive to phonological processing. Although less robust than the N400, the P200 has also been associated with general "mismatch". For example, in Hebrew, Barnea and Breznitz (1998) found reduced P200 amplitude for phonologically dissimilar relative to similar pairs in a target detection task and Hart, Perfetti, and Liu (reported in Liu, Perfetti, & Hart, 2003) found a similar result in English. Interestingly, Niznikiewicz and Squires (1996), found the opposite effect: greater negativity for homophones relative to orthographic controls at 200 ms. Regardless of direction of polarity these studies indicate that components that are sensitive to phonological processing may be early, thus comparison of both earlier and later components is necessary to examine potential differences in phonological processing.

### *1.3. Current study*

The goal of the current investigation was to test the hypothesis that the comprehension difficulties of less-skilled comprehenders may stem from a semantic processing difficulty (Nation & Snowling, 1998, 1999). To this end, we compared the ERP's of adult skilled and less-skilled comprehenders (matched on behavioral measures of decoding) during a semantic task (judging whether two words or pictures were related in meaning) and during a phonological task (judging whether two words had the same pronunciation).

To examine semantic processing further, we also tested processing of two types of semantic relationships: categorical (e.g., cat–horse) and categorical with association (e.g., cat–dog). Nation and Snowling (1999) found that poor comprehenders showed associative categorical priming but not pure categorical priming, suggesting that for poor comprehenders semantic priming may be driven by lexical co-occurrence and not by semantic activation per se. If this semantic vs. lexical co-occurrence difference is valid, we expected that less-skilled comprehenders' semantic ERP markers (e.g., N400) would differ primarily for categorical semantic processing. Furthermore, in order to determine if the semantic difficulty is limited to the access of semantic information from verbal labels, we examined ERP's in a picture comparison task. Although a few studies have looked at picture processing in poor comprehenders these studies involved matching pictures to words (Perfetti & Lesgold, 1979) or naming pictures (Nation et al., 2001), thus requiring explicit knowledge of verbal labels. To avoid overt verbal-semantic processing we chose to use a picture–picture task that required participants to judge the relationship between two pictures. Smaller N400's for semantically primed pictures relative to unprimed pictures have also been found in normal readers during picture processing (Hamm, Johnson, & Kirk, 2002). Thus, deviant N400 waveforms for less-skilled comprehenders (smaller reductions for related trials) during picture relatedness comparisons would suggest a semantic processing difficulty that extends beyond verbal semantic processing.

If our poor comprehenders' phonological processing/decoding skills are intact we predict that their ERP's during phonological comparisons should not differ from skilled comprehenders'. However, if poor comprehenders are also deficient in obtaining phonological codes from print, then their waveforms during our phonological task should differ from the waveforms of normal comprehenders'. In particular, if our less-skilled comprehenders have a phonological processing difficulty, we would not expect their N400 or P200 waveforms to differentiate between homophone and non-homophonic stimuli or we would expect them to differentiate less than the skilled-comprehenders.

In addition to providing direct assessments and comparisons of semantic and phonological processing by using ERP, testing adults in the current study will help to rule out a developmental lag or delay (assuming that basic development of the neural systems required for reading are fully developed by college in the normal population).

## 2. Methods

### 2.1. Participants

The original sample consisted of 39 adults who were all native English-speaking members of the University of Pittsburgh community. All participants were right handed and all participants reported normal or corrected to normal vision. Participants were compensated monetarily for their participation. Data from nine participants were excluded because these participants did not fit our criteria or because of EEG artifact in their recordings. Data from 30 participants (from 21 females and 9 males) was included in our analyses.

### 2.2. Skill assessments

The Nelson-Denny (ND) comprehension test was used to assess comprehension skill. To ensure that we were comparing readers who differed only on comprehension ability, participants were also given a timed test of non-word decoding ability, the Test of Word Reading Efficiency (TOWRE). Additionally, participants were given a non-verbal IQ test (Raven's Matrices) to ensure that there were no differences in general intelligence.

### 2.3. Materials

The experiment was divided into three blocks: the semantic-word task block, the semantic-picture task block, and the phonological task block. Each block used different stimuli. For the semantic word block, the stimuli were probe–target pairs that were either categorically related (lemon–pear), associatively and categorically related (cat–dog), or unrelated (bear–truck). Thirty-four pairs were categorically related, 34 pairs were associatively related, and 100 pairs were unrelated. Phrasal associates were avoided when possible (e.g., spider–web). Probes and targets were matched for frequency (Kucera & Francis, 1967 norms) across conditions and were normed for associative strength using the Edinburgh Associative Thesaurus (EAT) (Kiss, Armstrong, Milroy, & Piper, 1973) and Latent Semantic Analysis (LSA) (Landauer & Dumais, 1997) to ensure that categorically and associatively related pairs were in fact more highly associated than the categorically related pairs. (See Appendices A and C for stimuli and stimulus characteristics). All stimuli were shown only once and pairs were presented in random order.

For the semantic-picture task block, the picture stimuli consisted of pairs of related and unrelated picture pairs that were taken from the Snodgrass and Vanderwart (1980) normed line drawings. Related pictures shared category membership in one of the following categories: items of clothing, methods of transportation, animals, insects, body parts, furniture, vegetables or fruit. Unrelated pairs consisted of one member from one category and one member from a different category (no two items from overlapping categories such as vegetables and fruits or animals and insects were used to make an unrelated pair). One half of the picture pairs were semantically related (e.g., bear–tiger) and one half of the stimuli were unrelated (e.g., sheep–desk). Participants were not given any information about the types of categories before the experiment began. All stimuli were shown only once and pairs were presented in random order.

For the phonological task block, 54 homophone pairs and 54 orthographically similar non-homophone pairs were used. Words that made up homophone pairs and those that made up non-homophone pairs were not significantly different in length, frequency or letter overlap. All stimuli were shown only once and pairs were presented in random order. (See Appendices A and C for stimuli and characteristics)

### 2.4. Procedure

After ERP net application (see data acquisition and pre-processing), participants were seated in front of a Dell computer in a sound-attenuated room. For each trial in each block, participants saw a series of two pictures or words, presented one at a time and separated by a blank screen and were asked to decide if the two words or pictures they just saw were semantically related (in the first and second blocks) or if they had the same pronunciation (in the third block). Participants pressed the "1" key on the number key pad for a "yes" response and the "2" key on the number key pad for a "no" response. The display of the stimuli was controlled by E-Prime™ (Psychological Software Tools (PST), Pittsburgh, Pennsylvania). In all blocks during a given trial participants first saw a fixation cross (for 1000 ms), followed by the first stimulus (for 200 ms) then by a blank screen (for 200 ms), and finally by the second stimulus (which remained on screen until the participant responded). All word stimuli were displayed in black ink on a white background in 18 point Courier New font. For the word pairs, the first word appeared in all lower case letters and the second word (the target) appeared in all uppercase letters. Pictures were also presented in black on a white background and were all the same size (Fig. 1).

### 2.5. ERP data acquisition and pre-processing

All participants were fitted with Electro Geodesic Inc.'s (EGI's) Geodesic Sensor Net with a 128 Ag/AgCl electrode array. The potentials were sampled at a rate of 1000 Hz and

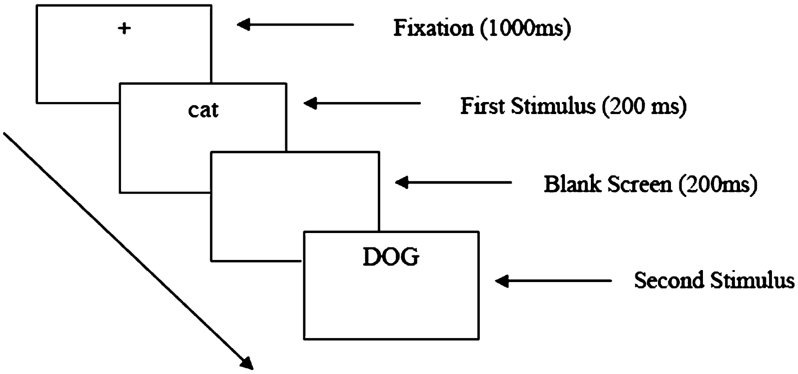

Fig. 1. Sample trial presentation for the semantic-word task.

were amplified 1000 times. Potentials were filtered with a 0.01-Hz high-pass hardware filter. Impedances were generally kept below a threshold of 40 kΩ with a max of 60 kΩ. A digital elliptical low pass filter of 30-Hz was applied and data were segmented into 1100 ms epochs starting 100 ms before the onset of the target word or picture. Bad channels were removed from the recordings and replaced with spherical spline interpolation from the remaining channels.[2] Segmented data were averaged across trials and edited for eye blinks and movement. Trials containing eye blink, eye movement or channel artifact were not included in the analysis.[3] As mentioned above, data from nine subjects in all were excluded; 30 remained in the final sample. After channel and subject exclusion, the data were re-referenced to the average of the 128 recording sites and were baseline corrected to the average activity during 100 ms before stimulus onset. Data were then combined for statistical extraction and analysis or grand averaged for examination of topographic maps and topographic plots. Only ERP's for correct responses were entered into the final analyses.

## 3. Results

Using a median split, the participants were divided into two skill groups of 15 skilled-comprehenders, 6 males and 9 females (ND Comprehension = 24.50 of 36 possible points; SD = 4.60) and 15 less-skilled comprehenders, 3 males and 12 females (ND Comprehension = 12.24; SD = 4.18), matched for decoding ability on the TOWRE (skilled comprehenders = 52.90 of 63 possible points; SD = 6.10; less-skilled comprehenders = 55.10; SD = 5.80), and for general intelligence on the Ravens Matrices (skilled comprehenders = 8.00 of 18 possible points; SD = 3.10; less-skilled comprehenders = 7.40; SD = 3.40).

### *3.1. Behavioral findings*

Accuracy and reaction time (RT) results were analyzed by separate, mixed, repeated measures analyses of variance (ANOVAs) with relatedness as a within subjects variable and skill as a between subjects variable.

#### *3.1.1. Accuracy*

Overall accuracy was high: all participants were better than 80% correct in all conditions. In the semantic word task, the effect of pair type was significant by subjects $F(2,56) = 33$, $p < .001$ and by items $F(2,161) = 6.28$, $p < .01$. Pair-wise comparisons revealed that both related conditions differed significantly from each other ($p < .01$) and from the unrelated condition ($p < .01$). Participants were most accurate for associatively related pairs (96.6%), less accurate for unrelated pairs (96.3%), and least accurate for categorically related pairs (89.0%).[4] There were no other main effects or interactions by subjects or by items.

In the semantic picture task, there were no main effects or interactions when analyzed by subjects. By items, there was a main effect of relatedness $F(1,64) = 5.06$, $p < .01$ participants were more accurate for related pairs (97%) than for unrelated pairs (94%). There were no other main effects or interactions by subjects or by items.

In the phonological task, the effect of relatedness was significant by subjects $F(1,28) = 8.8$, $p < .01$, but not by items $p > .1$. Participants showed higher accuracy for related (97%) compared with unrelated pairs (95.6%). There

---

[2] Overall very few bad channels were identified. Most were on the face or periphery and thus were not part of the head electrodes used in our analyses. For those channels used in our primary analysis, one participant in the low skill group had a spherical spline interpolation for channel F3 (25). Additionally the P4 (87) channel was replaced by spherical spline for a few trials <5% in the semantic and <2% in the phonological condition for one low-skill participant. Finally three participants (two less skilled and one skilled) had spherical spline interpolations for <2% of their trials total for channel f4 (124).

[3] Average number of good trials in the final analysis: phonological task (averaged across related and unrelated trials: $M = 45.4$, SD = 7.1); semantic word task related trials (averaged across associatively related and categorically related: $M = 27.0$, SD = 4.1); semantic-word unrelated trails ($M = 82.6$, SD = 15.9); semantic-picture trials (averaged across related and unrelated: $M = 28.4$, SD = 4.3).

[4] This trend for poorer performance in the categorically related condition reflects some ambiguity in this decision, relative to the associatively related condition given that speed was stressed (in particular, three pairs had especially low accuracy and thus were presumably less transparent to students: dolphin-human, drought-blizzard and almond-peanut). Only accurate trials were used for the RT and ERP analysis.

were no other main effects or interactions by subjects or by items.

### 3.1.2. Reaction time

Only correct trials were considered in the reaction time analysis.

In the semantic-word task, the effect of relatedness was significant by subjects $F(2,56) = 40.0$, $p < .001$ and by items $F(2,156) = 20.86$, $p < .001$. Pair-wise comparisons revealed that participants responded faster to associatively related pairs than categorically related pairs $p < .05$ or unrelated pairs $p < .021$. Categorically related and unrelated pairs did not differ ($p > .1$). The effect of skill was significant by subjects $F(1,28) = 432$, $p < .001$ and by items $F(1,156) = 132.8$, $p < .001$. Skilled-comprehenders were faster overall at making the semantic judgments than less-skilled comprehenders. There was also a significant interaction between skill and relatedness by items $F(2,156) = 9.4$, $p < .001$, but this effect was not significant by subjects. There were no other main effects or interactions by subjects or by items.

In the semantic-picture task, the effect of skill was significant by subjects $F(1,28) = 392.0$, $p < .001$ and by items $F(1,64) = 40.7$, $p < .001$. Skilled comprehenders were faster overall than less-skilled comprehenders. There were no other main effects or interactions by subjects or by items.

In the phonological task, the effect of relatedness was significant by subjects $F(1,28) = 5.29$, $p < .05$ and by items $F(1,104) = 6.04$, $p < .05$. Participants made faster responses to related pairs than to unrelated pairs. The effect of comprehension skill was significant by subjects $F(1,28) = 4.36$, $p < .05$ and by items $F(1,104) = 142.0$, $p < .001$. Less-skilled comprehenders were slower overall to judge homophony than skilled comprehenders. The interaction between relatedness and skill was also significant by items $F(1,104) = 4.425$, $p < .05$, but not by subjects. There were no other main effects or interactions by subjects or by items.

Table 1 contains the full set of means, standard deviations, and ranges of reaction times for all tasks.

Table 1
Reaction time means, ranges and standard deviations for skilled and less-skilled comprehenders in each condition

| Condition | Min | Max | Mean | Std. Dev |
|---|---|---|---|---|
| *Skilled comprehenders (N = 15)* | | | | |
| Associatively related | 453.71 | 835.00 | 581.89 | 132.65 |
| Categorically related | 526.33 | 1025.50 | 680.37 | 165.43 |
| Unrelated | 473.31 | 974.34 | 660.23 | 171.92 |
| Homophone | 397.17 | 867.30 | 560.99 | 132.97 |
| Non-homophone | 421.25 | 935.12 | 581.93 | 138.60 |
| Related picture | 376.91 | 946.61 | 566.19 | 149.77 |
| Unrelated picture | 421.27 | 1138.77 | 594.26 | 185.09 |
| *Less-skilled comprehenders (N = 15)* | | | | |
| Associatively related | 453.52 | 1011.64 | 648.43 | 171.62 |
| Categorically related | 592.71 | 1322.39 | 763.69 | 203.01 |
| Unrelated | 544.61 | 1404.95 | 787.16 | 250.06 |
| Homophone | 390.71 | 926.73 | 654.51 | 135.86 |
| Non-homophone | 473.75 | 1124.13 | 696.28 | 157.41 |
| Related picture | 460.25 | 1179.66 | 648.71 | 180.05 |
| Unrelated picture | 467.47 | 1188.74 | 706.15 | 213.27 |

### 3.2. Electrophysiological findings

For initial inspection, we viewed 128-channel waveform plots of the grand-averaged files for skilled comprehenders and for the less-skilled comprehenders. We compared ERP response to associatively related, categorically related, and unrelated targets; we then compared related and unrelated picture targets; and finally we compared waveform plots for homophonic vs. non-homophonic targets. We also viewed "Student's *t*-test" topographic maps[5] ($\alpha = .05$) for each condition within each task compared to each other condition in that particular task. Time points and electrodes were selected for statistical testing based on previous research on N400 and P200, and confirmed with visual inspection of waveform plots and on the results from the Student's *t*-test topographic maps (see Frishkoff et al., 2004 for additional reports using EGI's t-test maps). Two time windows were sensitive to group and/or condition differences and were thus subjected to further analysis: 150–250 ms (P200) time window for the two word tasks and a 300–600 ms (N400) time window for all tasks. One other very early window (N100) showed relatedness effects in the picture task, most likely a visual processing difference due to the fact that the related and unrelated pictures were not controlled for visual similarity. Because of this visual similarity difference this effect is not discussed further here (see Landi, 2006, for a complete discussion).

*10–20 Analysis.* For statistical testing of the P200 effect, average activations between 150 and 250 ms for the nine electrodes of the international 10–20 system (F3, Fz, F4, C3, CZ, C4, P3, PZ, P4) were analyzed using a repeated–measures, mixed ANOVA with relatedness (related [two levels for the semantic word task], unrelated), lobe (frontal, central, parietal), and hemisphere (right, left, midline) as within subjects variables and skill (less-skilled comprehenders, skilled comprehenders) as a between subjects variable. Based on previous research localizing the N400 component to the central-posterior parietal region we selected the posterior six electrodes of the international 10–20 system (C3, CZ, C4, P3, PZ, P4)[6] for statistical testing of the N400. Average amplitudes between 300 and 600 ms were analyzed using an ANOVA of the form described above. Interactions were further subjected to pair-wise *t*-tests. Latency analyses were also conducted where appropriate to determine if

---

[5] The Student's *t*-test tool is available in the newest beta version of Netstation. Student's *t*-test topographic maps have appeared in published work by Frishkoff, Tucker, Davey, and Scherg (2004).

[6] These electrodes were selected because they encompass the strongest site of the N400 effect and because the N400 is typically observed in these sites. Furthermore, exclusion of the frontal three electrodes ensured that we were analyzing the N400 and not the MFN, a more anterior component, which although similar in shape in time course to the N400 is thought to have a different neural generator and is linked to attentional processes (see Frishkoff et al., 2004; Landi, 2006).

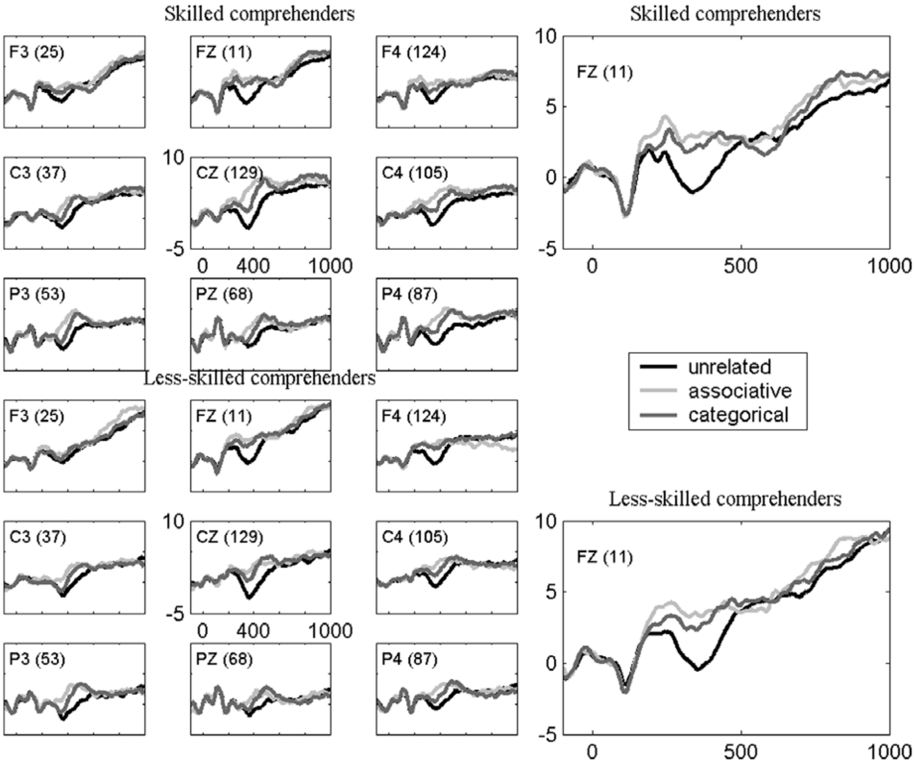

Fig. 2. ERP response to unrelated (solid black), categorically related (gray) and categorically and associatively related (dashed) targets, for skilled comprehenders and less-skilled comprehenders in the semantic-word task. The FZ electrode is enlarged for viewing of the P200.

component peaks differed in time across condition or between groups.

*Cluster analysis.* In order to take advantage of the high density recording, and to provide a confirmation of the stability of our initial electrode selection, additional comparisons were run on 9 electrode clusters for the P200 analyses (right-frontal: electrodes #25, 21, 13, 30, and 29; medial-frontal: electrodes #11, 4, 20, 12, 6, and 5; left-frontal: electrodes #124, 119, 118, 113, and 112; left-central: electrodes #42, 36, 37, 43, 41, 47, and 48; medial-central : electrodes #129, 81, 55, 32, 7, and 107; right-central: electrodes #104, 105, 11, 110, 94, 99, and 103; left-parietal: electrodes #53, 52, 60, 66, 59, and 51; medial-parietal: electrodes #62, 68, 61, 79, 67, 73, and 78; left-parietal: electrodes #87, 93, 98, 86, 92, and 85) and six central/posterior electrode clusters for the N400 (left-central, medial-central, right-central, left-parietal, medial-parietal, left-parietal). Clusters were selected to provide full head coverage in a manner consistent with the 10–20 system; thus, each cluster represents an electrode of the 10–20 system and 4–6 adjacent electrodes. This method for electrode cluster selection is similar to that of Perfetti, Wlotko, and Hart (2005).[7] Average amplitudes across all electrodes in a cluster for the relevant time period (P200 or N400) were entered in as separate factors in a mixed, repeated measures ANOVA. Main effects and interactions are reported in the main text, pair-wise comparisons for effects involving skill are reported in Appendix B.

### 3.2.1. Semantic-word task

*P200.* The initial detection of semantic incongruity was a positive going waveform that began around 150 ms and continued through 250 ms, just before the beginning of the N400 separation (Fig. 2). This early sensitivity to semantic incongruity was somewhat unexpected given the concentration in the literature on N400 effects for semantic tasks. However, studies investigating the recognition potential (RP) component which peaks at 200 ms, have detected such early sensitivity to semantic manipulations (Martin-Loeches, Hinojosa, Gomez-Jarabo, & Rubia, 2001) and other studies have found modulation of the P200 for semantic associates (Coulson, Federmeier, Van Petten, Kutas, & Memory & Cognition, 2005), thus suggesting that semantic processing differences may be detectable by 200 ms in some paradigms. In our study, this early difference was larger for skilled comprehenders (see below). If only skilled comprehenders are sensitive to semantic relationships this early in processing, previous studies may have failed to detect this P200 difference because they used average readers (a mix of skilled and less-skilled comprehenders not separated by skill).

The ANOVA revealed a main effect of relatedness $F(2,56)=10.0$, $p<.001$, Pair-wise comparisons revealed that the associatively related condition response was significantly different from the categorically related condition response $p<.01$, and that the associatively related response

---

[7] Unlike Perfetti et al. (2005), in our clusters the electrodes are not all adjacent to the 10–20 electrode because we wanted to avoid inclusion of a particular electrode in multiple clusters but electrodes are all adjacent to each other.

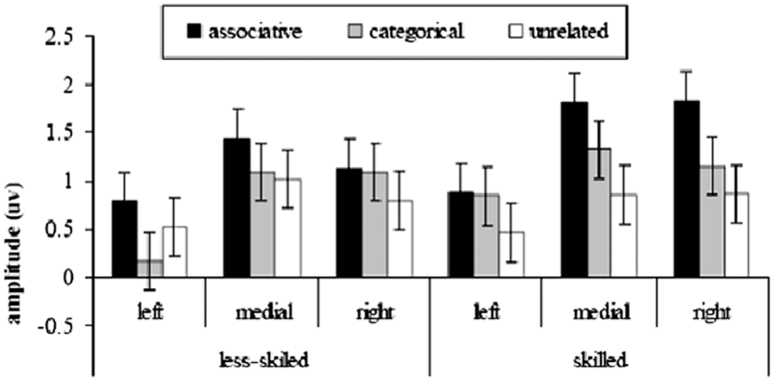

Fig. 3. Average amplitude of the ERP response during the semantic–word task, during the 150–200 ms (P200) time window, collapsed across the nine electrodes of the 10–20 system, comparing relatedness conditions (categorical and associative, associative, unrelated) and hemisphere (left, medial, right) for skilled and less-skilled comprehenders.

was significantly different from the unrelated response, $p < .01$; however, the difference between the unrelated response and the categorically related response was only marginal $p = .08$. Furthermore, there was a significant relatedness by lobe interaction $F(4, 112) = 4.0$, $p < .01$, and a significant relatedness by lobe by hemisphere interaction $F(8, 224) = 3.4$, $p < .01$, suggesting that there was greater P200 sensitivity to semantic relationship in the frontal lobe and over the medial and right hemispheres. Pair-wise comparisons for associative–unrelated conditions were significant in left, medial and right frontal regions, all $p < .01$ and medial and right central regions, all $p < .01$. Comparisons for central left and all parietal regions were non significant, all $p > .1$. For associative–categorical, significant differences were found in all frontal electrodes, all $p < .05$, and the right central electrode, $p < .05$. For categorical–unrelated the medial-frontal comparison was significant $p < .05$ and there were two marginal differences, one in right frontal $p = .054$ and one in right parietal, $p = .059$, the other comparisons were all non significant, all $p > .1$. Importantly, there was also a significant comprehension skill by hemisphere by relatedness interaction $F(4, 112) = 5.0$, $p < .01$.Although both groups showed reductions for related pairs, less-skilled comprehenders showed smaller P200 reductions for related pairs (particularly in the medial and right hemispheres) than skilled comprehenders. Furthermore, less-skilled readers showed only one significant difference between any of the relatedness conditions (between associatively and categorically related conditions) whereas skilled readers showed significant differences in four of the nine pair-wise comparisons (see Table 2 for mean differences and $p$ values). There were no other main effects or interactions for the skill or relatedness variables.[8]

Peak latency analyses for the P200 effect in this task revealed no differences in latency between the two skill groups (Fig. 3).

Table 2
Average amplitude differences during the 150–250 ms time window (P200) for left, medial, and right hemispheres, for skilled and less-skilled comprehenders for the semantic-word task

| Hemisphere | Associative–categorical | Associative–unrelated | Categorical–unrelated |
|---|---|---|---|
| *Skilled comprehenders (N = 15)* | | | |
| Left | .03 | .42 | .38 |
| edial | .49 | .96** | .47* |
| Right | .68* | .97*** | .28 |
| *Less-skilled comprehenders (N = 15)* | | | |
| Left | .62** | .27 | .34 |
| Medial | .35 | .42 | .06 |
| Right | .04 | .33 | .28 |

* $p < .05$.
** $p < .01$.
*** $p < .001$.

Cluster analyses confirmed the general pattern of results from the original ANOVA: there was a main effect of relatedness $F(2, 56) = 5.29$, $p < .01$, with most negative trajectories for unrelated words, intermediate trajectories for categorically related words, and most positive trajectories for associatively related words. There was also a relatedness by cluster region interaction $F(16, 448) = 9.67$, $p < .001$, with larger P200 effects in right-frontal regions. Critically, there was also a relatedness by skill by cluster interaction $F(16, 442) = 1.86$, $p < .05$, with generally larger differences between relatedness conditions for the skilled readers in the frontal and central clusters. Mean differences and $p$ values from pair-wise comparisons are shown in Table A2.

*N400.* There was a large negative-going waveform that began anteriorly around 250 ms and continued to grow and move posteriorly, until around 600 ms (Figs. 2 and 4). This N400 component was larger (more negative) for unrelated compared to related targets. Furthermore, the N400 was larger for categorically related targets than for associatively related targets. As Figs. 2 and 4 show, these effects were observable for less-skilled comprehenders but were smaller than the effects for skilled-comprehenders.

The ANOVA revealed a main effect of relatedness $F(2, 56) = 46.0$, $p < .001$. Pair-wise comparisons confirmed

---

[8] Main effects or interactions not associated with skill or relatedness (e.g., main effect of hemisphere) are not reported here.

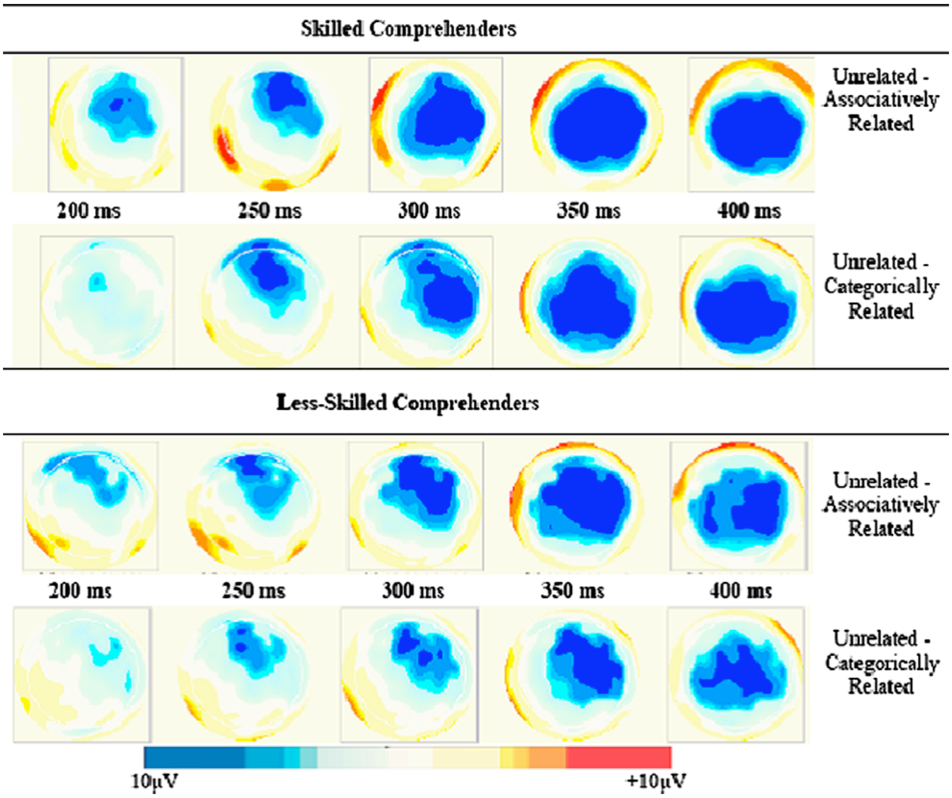

Fig. 4. Student's *t*-test maps for unrelated minus associative and unrelated minus categorically related conditions for skilled and less-skilled comprehenders. Darker shades correspond to greater amplitude differences (blue for negative differences and red for positive differences).

that each relatedness condition was different from each other relatedness condition (unrelated–associatively related, $p < .01$; unrelated–categorically related $p < .01$; categorically related–associatively related $p < .01$). Furthermore, there was a significant relatedness by lobe by hemisphere interaction $F(4,112) = 15.26$, $p < .01$, indicating a larger N400 effect in the central lobe and middle hemisphere. Pair-wise comparisons of associated - unrelated conditions were significant in all lobes and hemispheres, all $p < .01$. For associative–categorical, left central, medial central, right central and left parietal regions were significant at $p < .01$, medial parietal and right parietal were not significant $p > .05$, although right parietal was marginal $p = .07$. For categorically related–unrelated all pair-wise comparisons were significant, all $p < .01$. Importantly, there was also a relatedness by comprehension skill by hemisphere interaction $F(4,112) = 2.7$, $p < .05$. Less-skilled comprehenders showed smaller N400 reductions than skilled comprehenders for categorically related targets and for associatively related targets, particularly in the right and medial hemispheres. Unlike skilled comprehenders, less-skilled comprehenders did not show any additional N400 reduction for associatively related pairs relative to categorically related pairs. (Table 3 shows pair-wise mean differences and $p$ values). Differences between the skill groups in the size of the N400 can also be clearly seen by in the topographic maps (Fig. 4). These findings suggest that although both groups are sensitive to semantic relationship difference, skilled comprehenders are more sensitive to both types of semantic relationships (categorical and associative) compared to less-skilled comprehenders when sufficiently sensitive measures are used Fig. 5

Table 3
Average amplitude differences during the 300–600 ms time window (N400), between relatedness conditions for left, medial, and right hemispheres, for skilled and less-skilled comprehenders for the semantic-word task

| Hemisphere | Associative–categorical | Associative–unrelated | Categorical–unrelated |
|---|---|---|---|
| *Skilled comprehenders (N = 15)* | | | |
| Left | 0.88*** | 2.2*** | 1.37*** |
| Medial | 1.0* | 2.92*** | 1.92** |
| Right | 1.0* | 3.0*** | 1.9*** |
| *Less-skilled comprehenders (N = 15)* | | | |
| Left | 0.78 | 2.2*** | 1.4** |
| Medial | 0.09 | 1.9*** | 1.89 |
| Right | 0.42 | 1.8*** | 1.4*** |

* $p < .05$.
** $p < .01$.
*** $p < .001$.

Peak latency analyses for the N400 effect in this task revealed no differences in latency between the two skill groups.

Cluster analyses confirmed the general pattern of results from the original ANOVA: Main effect of relatedness $F(2,56) = 31.95$, $p < .001$, with the largest reductions for the associated pairs, followed by the categorically related pairs.

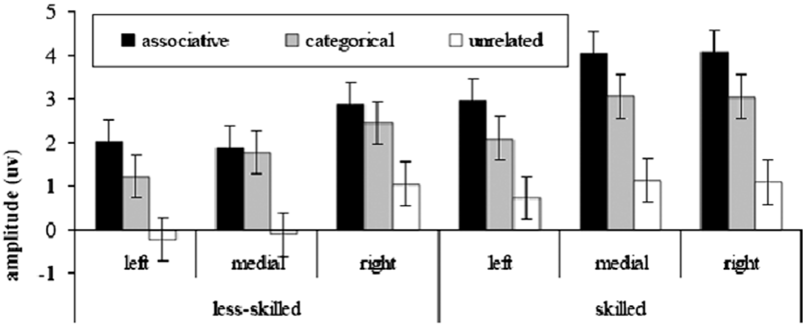

Fig. 5. Average amplitude of the ERP response during the semantic–word task, during the 300–600 ms (N400) time window, collapsed across the posterior six electrodes of the 10–20 system, comparing relatedness conditions (categorical and associative, associative, unrelated) and hemisphere (left, medial, right) for skilled and less-skilled comprehenders.

There was also a relatedness by cluster region interaction $F(10,280)=7.39$, $p<.01$, with the largest N400 reductions in central regions. And, importantly, relatedness by skill by cluster $F(10,280)=1.98$, $p<.05$, with generally larger differences between skill groups in right parietal regions. Mean differences and $p$ values from pair-wise comparisons are shown in Table A3.

### 3.2.2. Semantic picture task

*N400*. Fig. 6 shows a large N400 response beginning at 300 ms and continuing until 600 ms with more negative-going waveforms to unrelated pictures than to related pictures. The N400 for picture mismatches was similar in shape to previous reports (Federmeier & Kutas, 2001; Hamm et al., 2002). This N400 effect was largest in centro-parietal regions (although this difference is less clear for pictures than for words). The ANOVA revealed a main effect of relatedness $F(1,28)=7.4$, $p<.05$, with more positive ERP responses to related picture targets than to unrelated picture targets. There was also a relatedness by lobe interaction $F(1,28)=11.4$, $p<.01$, and a relatedness by lobe by hemisphere interaction $F(1,56)=6.95$, $p<.01$. The N400 effect was larger in central than in parietal regions and over the medial rather than the left or right hemispheres. Pair-wise comparisons revealed that related amplitudes were more positive than unrelated amplitudes in the central lobe, left hemisphere $p<.01$, the central lobe, medial hemisphere $p<.01$, the central lobe, right hemisphere $p>.01$ and the left hemisphere of the parietal lobe $p<.05$. Amplitudes between the two conditions did not differ in the other two hemispheres of the parietal lobe, all $p>.05$. There were no interactions with comprehension skill and there were no differences in peak latency between the groups.

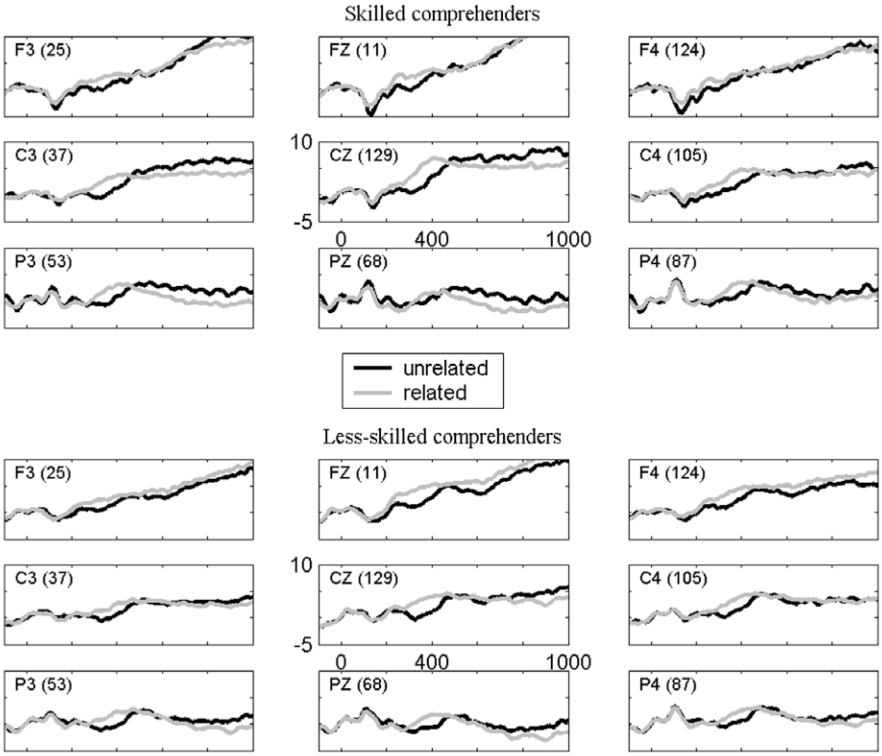

Fig. 6. ERP response to unrelated (solid black) and related (dashed) targets for skilled and less-skilled comprehenders, for the semantic-picture task.

Cluster analyses confirmed the general pattern of results from the original ANOVA: There was no main effect of relatedness $F(2,26)=2.2$, $p=.13$, but there was a relatedness by cluster region interaction $F(10,130)=3.86$, $p<.001$, confirming that the picture N400 was larger in the medial-central region. There were no main effects or interactions with relatedness and comprehension skill.

### 3.2.3. Phonological task

*P200.* As in the semantic word task, the phonological task produced a waveform beginning around 150 ms and continuing until 250 ms that was more pronounced in anterior regions (Fig. 7).

The ANOVA revealed a main effect of relatedness type $F(1,28)=8.1$, $p<.01$, with non-homophonic targets generally more negative than homophonic targets. There was also a relatedness by lobe by hemisphere interaction $F(4,112)=5.8$, $p<.01$, with the largest reduction for homophone targets occurring in the frontal and central regions and in the left hemisphere. Pair-wise comparisons revealed significant differences between homophone and non-homophone conditions in left frontal $p<.01$, medial frontal $p<.01$, left central $p<.05$, medial central $p<.05$ and right central $p<.05$. Differences in right frontal, left parietal, central parietal and right parietal were not significant, all $p>.05$. There were no interactions with relatedness or comprehension skill.

Due to a lack of true peak (see Fig. 7) latency analyses could not conducted for this component in this task.

Cluster analyses confirmed the general pattern of results from the original ANOVA: There was a main effect of relatedness $F(1,28)=6.7$, $p<.05$ and a relatedness by cluster region interaction $F(8,224)=5.39$, $p<.001$. The larger reductions were seen in the frontal and central regions and in the left and medial hemispheres. There were no main effects of or interactions with comprehension skill.

*N400.* As with the two semantic tasks, the phonological task produced a N400 waveform that began anteriorly around 250 ms and moved posteriorly until 600 ms (Fig. 7). The ANOVA examining the N400 effect revealed a main effect of relatedness $F(1,28)=69.0$, $p<.001$, with non-homophonic targets more negative than homophonic targets. There was also a relatedness by hemisphere interaction $F(2,56)=28.3$, $p<.001$, and a relatedness by lobe by hemisphere interaction $F(2,56)=17.4$, $p<.001$. Pair-wise comparisons revealed significant differences between homophone and non-homophone conditions in all lobes and hemispheres: left central $p<.001$, medial central $p<.001$, right central, $p<.001$, left parietal, $p<.001$, medial parietal, $p<.001$ and right parietal $p<.001$ but the mean differences were largest in parietal lobe and medial hemisphere. There were no significant interactions with comprehension skill for the phonological task and there were no effects of peak latency between the skill groups.

Cluster analyses confirmed the general pattern of results from the original ANOVA: there was a main effect of relatedness $F(1,28)=54.7$, $p<.001$ and a relatedness by cluster region interaction $F(5,140)=10.69$, $p<.001$. Larger reductions were seen in the parietal lobe and medial hemisphere. There were no main effects of or interactions with comprehension skill.

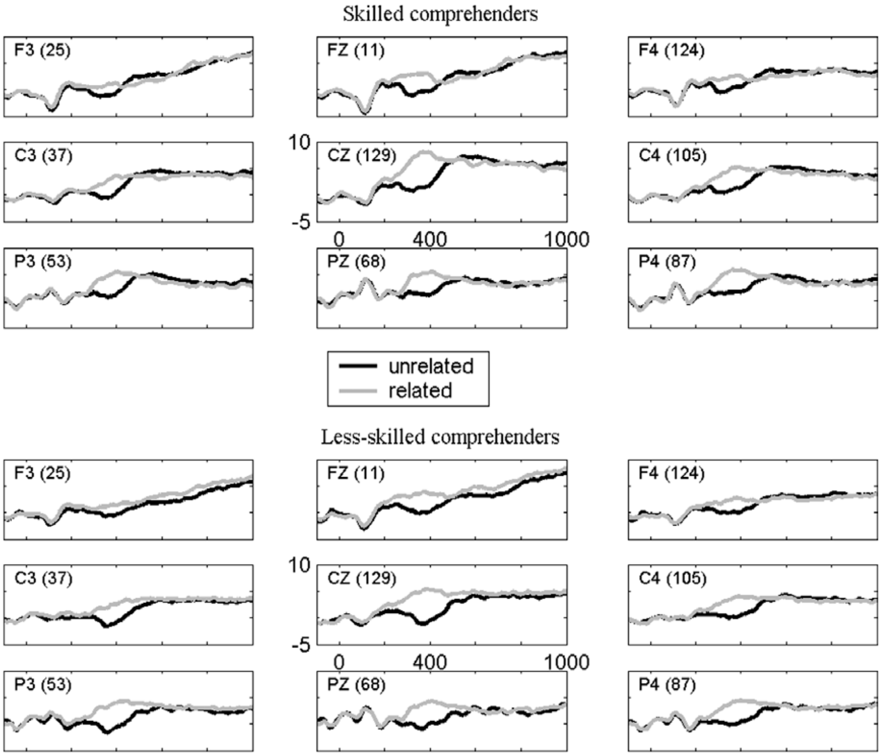

Fig. 7. ERP response to unrelated (solid black) and related (dashed) targets for skilled and less-skilled comprehenders, for the phonological task.

### 3.2.4. Post hoc RT covariate analysis

Participants' responses sometimes occurred within the N400 window and this was more common for the skilled comprehenders than the less-skilled comprehenders because skilled comprehenders were faster than less-skilled comprehenders in all tasks. If ERP's reflect response-related components (e.g., an off-potential associated with the stimulus disappearing from the screen after the response), then our skill-related ERP effects could be due to RT differences, not underlying processing differences. To test this possibility, we conducted mixed, repeated measures ANOVA's with RT added as a covariate for each task for both the N400 and N200 effects. To examine the effects of RT on skill, we entered average subject RT (averaged across condition) separately for each task. These analyses used the same electrodes and time windows described for the 10–20 analyses. These additional analyses examined whether our skill difference findings were driven by differences in RT. As in the analyses reported above, in the semantic word task there was a significant relatedness by hemisphere by skill interaction for the P200 $F(4,108) = 4.52$, $p < .001$ and for the N400 $F(4,108) = 2.85$, $p < .05$, but there were no significant main effects of skill or interactions with skill and relatedness in either the semantic picture task or the phonological task ( all $p > 0.05$). That is, even when individual differences in RT were covaried out, there were still comprehension skill differences in the semantic word task N400 and P200, but not in the semantic picture task N400 or phonological task P200 or N400.

## 4. Summary and discussion

The largest differences between skill groups in our experiment were seen in the semantic-word task, which elicited a P200 and an N400 that were sensitive to semantic incongruity. Furthermore, these differences remained significant after the RT covariate was added and when the data were subjected to the larger scale cluster analysis, suggesting that these effects are stable and not driven by differences in RT.

The P200 component was larger (more positive) for semantically related pairs relative to semantically unrelated pairs and was larger for associatively related pairs than for categorically related pairs. Skilled comprehenders showed larger P200 differences between related and unrelated pairs than less-skilled comprehenders. Some researchers have suggested that such early semantic effects at the P200 may be due to the onset of an N400 (Coulson et al., 2005). Although this hypothesis is consistent with the fact that both components are sensitive to skill differences in semantic but not phonological processing, the locus of the effects was somewhat different (the P200 was significantly more frontal than the N400), suggesting a different source. As this study was not designed to asses this relationship, further experimentation is required to determine the precise relationship between these two components.

The N400 was also larger (more negative) for semantically unrelated relative to semantically related pairs and it was more negative for categorical than for associative pairs. Less-skilled comprehenders showed smaller categorical N400 and associative N400 reductions relative to unrelated pairs, compared with skilled comprehenders, suggesting differences in semantic processing between the skill groups for both types of semantic relationships. Less-skilled comprehenders also differed from skilled comprehenders in that they failed to show any additional N400 reductions for associatively related pairs beyond their reductions for categorical pairs. Our finding of differences in both categorical and associative processing in an ERP task suggests that less-skilled readers differ from skilled readers in processing of both associative and categorical relationships. We emphasize, however, that both groups showed sensitivity to semantic relationships, so the differences between skill groups in the ERP were a matter of degree. By using a measure sufficiently sensitive to detect both categorical and associative relationships, our finding extends previous behavioral studies with children that found differences only for categorical relationships (Nation & Snowling, 1999). ERP waveforms picked up differences between skill groups in the semantic word task that were not observable in the RT data: Skilled and less skilled readers showed similar patterns of facilitation in RT (both showed associative but not categorical facilitation) but significant differences between their N400 response (greater N400 reductions for skilled readers). This pattern is consistent with other studies that have found differences in ERP when behavioral measures failed to show differences (e.g., Kiefer & Brendel, 2006). These differences are often attributed to the fact that ERP and RT might measure somewhat different aspects of processing. For example, RT measures processes that are related to a response (e.g., decision processes) and those that vary as a function of cognition whereas the N400 may reflect only those that vary as a function of cognition (e.g., automatic spreading activation). For the purpose of the current study, we can conclude that our differences in N400 during semantic word processing for categorical and associative relationships were caused by a cognitive process that is not always detected in RT.

Although we found N400 and P200 skill differences in the semantic-word task, we found no skill differences in the semantic-picture task, suggesting that word-semantic processes rather than general semantic processes were the source of the skill differences. However, the picture task was somewhat easier for participants than the semantic-word task (they were faster and more accurate in the semantic-picture task), possibly due to the fact that participants may have been able to use visual similarity cues (as these were not matched across related and unrelated trials). Further examination utilizing a more difficult visual-semantic processing task would provide a more definitive test of this verbal/visual-semantic distinction.

Consistent with the assumption that we identified subjects who had only a semantic level and not a decoding level problem, there were no differences between skill groups in the P200 component or N400 components during the phonological task. Less-skilled comprehenders did differ from skilled comprehenders in their overall processing time in the phonological task (as they did in the semantic task). Specifically, less-skilled comprehenders were slower to make the sound-alike judgments than skilled comprehenders (they were also slower to make the semantic relatedness judgment), which suggested some type of processing difference between the two groups in the phonological task that did not show up in the ERP record. This difference between the ERP and behavioral data, as we noted above, may suggest that the two groups differ at a later decision-making stage or in a more general processing capacity that is not reflected in ERP's. We cannot rule out the possibility that slower decisions reflect less certainty about phonological properties of words, which could be reflected in a decision checking process.

Although no specific late component differences were identified in the initial analysis that might reflect such a decision checking process, the overall shape of the waveforms differed between skill groups in all tasks. The skilled-comprehenders tended to have greater positive deflection overall, regardless of manipulation, and the less-skilled comprehenders tended have more negative deflection (see Figs. 2, 6 and 7). There are relatively few ERP studies comparing skilled and less skilled readers, making this pattern difficult to interpret; however, it is possible that this general difference in amplitude could reflect an underlying processing difference. One study (Meyler & Breznitz, 2005) using a mastoid reference reported a related difference: dyslexic readers showed more negative amplitudes than controls. Furthermore, close examination of figures from other ERP studies that examined reading skill differences (although not limited to comprehension differences per se), revealed that less-skilled readers had generally more negative waveforms (Perfetti et al., 2005; Yang, Perfetti, & Schmalhofer, 2005 (both using averaged referenced ERP's); Segalowitz, Wagner, & Menna, 1992 (linked ear referenced ERP's)). This examination suggests that this polarity difference may be real, and that it deserves further investigation.

Our findings are generally consistent with the hypothesis that some adult less skilled readers differ from skilled readers in semantic processing without corresponding differences in phonological processing. The question becomes how do less skilled comprehenders come to have a semantic processing weakness? The possibilities include some type of congenital deficit in an individual's ability to acquire or access semantic knowledge and, alternatively, a lack of *relevant* experience required to develop necessary lexical-semantic skills. Although there are likely to be some innate differences between individuals in both the ability to acquire and the ability to access word meanings, the primary way in which semantic representations are built is through co-occurrence of information, suggesting the most likely cause of semantic difficulty in this population is a failure to build appropriate representations. MacDonald and Christiansen (2002) demonstrated how lack of experience might affect comprehension in sentence processing and Sandak et al. (2004) demonstrated that training on semantic features improves word learning. Hart (2005) showed individual differences in the rate of learning "meanings" in a novel artificial language that were then manifest in comprehension performance using the novel language. Furthermore, existence proofs provided by connectionist models demonstrate that robust representations can be built up through multiple encounters with both a word's phonological and semantic characteristics (Harm & Seidenberg, 2004; Plaut, McClelland, Seidenberg, & Patterson, 1996). Of course, the relationships among experience, semantic knowledge and comprehension skill are likely to be reciprocal. Fewer high quality experiences lead to impoverished semantic representations which lead to poor comprehension ability, which discourages reading, which furthers the trend for inadequate experiences. Further research involving longitudinal examinations and training paradigms will be important for elucidating the exact nature of these differences in individuals with comprehension difficulty.

## 5. Conclusion

The present study demonstrated differences between skilled and less-skilled comprehenders in a semantic processing task that are not likely to be caused by differences in decoding skill or by a developmental lag. Our findings suggest that semantic processing ability may be an important underlying factor in reading comprehension skill regardless of age or developmental status. These findings are consistent with behavioral studies of children with specific comprehension impairment that have found differences in semantic processing ability in a variety of behavioral tasks (Nation & Snowling, 1998, 1999; Nation, Snowling, & Clarke, 2005) and with ERP studies of adult less-skilled comprehenders that have found these readers to be slow to access and integrate meaning when an inference was required for text comprehension (Yang et al., 2005).

We did not find differences in ERP waveforms between the two groups in a phonological task, confirming that our skill groups did not differ in phonological processing ability. The groups did differ in RT for both the semantic and phonological task, which implies an additional processing difference, possibly in decision making. The lack of skill differences in the semantic-picture task indicates that differences in semantic processing between skilled and less-skilled comprehenders were primarily limited to the verbal-semantic domain (with the caveat that our picture task may have been easier than our semantic word task, which could mask possible underlying differences). Our finding of greater differences in semantic processing than in phonological processing in poor comprehenders is generally consistent with the semantic deficit hypothesis (Nation & Snowling, 1998). These findings are consistent also with the lexical quality hypothesis (Perfetti & Hart, 2001), which posits the importance of adequate processing of all lexical-level components (semantics, phonology and orthography) for successful reading comprehension. We speculate that differences in relevant reading experience between the two groups are a likely cause of the observed differences in semantic processing.

## Acknowledgments

We thank Drs. Kenneth Pugh, Stephen Frost and Daniel Mirman for their helpful comments and editorial assistance. This research was supported in part by the Institute of Education Sciences, Award No. R305G020006A: "Word learning and comprehension: New laboratory approaches and classroom studies", September 2002–2005.

## Appendix A

The following table provides additional information relevant to reproduction of the findings reported in this paper Table A1.

## Appendix B. Supplemental analysis

See Tables A2 and A3.

## Appendix C

Word pairs used in the semantic-word and phonological tasks:

*Associatively and categorically related pairs*: red–GREEN, diamond–RING, hugs–KISSES, sugar–SPICE, girls–BOYS, brother–SISTER, dog–CAT, king–QUEEN, moon–STAR, salt–PEPPER, boat–SHIP, coat–HAT, comb–BRUSH, cup–SAUCER, table–CHAIR, pencil–PEN, mother–FATHER, computer–MOUSE, car–TRUCK, orange–APPLE, black–WHITE, doctor–NURSE, school–TEACHER, snow–ICE, spaghetti–MEATBALLS, lion–TIGER, pillow–SLEEP, bread–BUTTER, cottage–CHEESE, cap–GOWN, suit–TIE, coffee–CREAM, toaster–OVEN, circle–SQUARE.

*Categorically related pairs*: mug, GLASS, vein–ARTERY, clam–SHRIMP, steel–COPPER, chicken–HORSE, green–PINK, lake–MOUNTAIN, nose–HEAD, button–ZIPPER, airplane–TRAIN, kite–BALLOON, bed–DESK, book–MAGAZINE, sweater–SKIRT, violin–GUITAR, roof–WALL, banana–TOMATO, human–DOLPHIN, shirt–SOCK, stomach–HEART, church–TEMPLE, swim–RUN, stereo–TELEVISION, necklace–RING, spider–FLY, peanut–ALMOND, physics–BIOLOGY, blizzard–DRAUGHT, sage–NUTMEG, velvet–COTTON, rose–LILY, penny–QUARTER, cardinal–ROBIN, tango–WALTZ.

Table A2

Average amplitude differences during the 150–250 ms time window (P200) between relatedness conditions for nine electrode clusters for skilled and less-skilled comprehenders for the semantic-word task

| Cluster | Associative–categorical | Associative–unrelated | Categorical–unrelated |
|---|---|---|---|
| *Skilled comprehenders* | | | |
| Frontal | | | |
| Left | 0.39 | 0.90$^{**}$ | 0.51$^{*}$ |
| Medial | 1.04$^{**}$ | 1.57$^{****}$ | 0.53$^{*}$ |
| Right | 1.08$^{**}$ | 1.32$^{****}$ | 0.24 |
| Central | | | |
| Left | −0.27 | 0.01 | 0.28 |
| Medial | 0.65 | 1.18$^{**}$ | 0.53$^{*}$ |
| Right | 0.37 | 0.60$^{*}$ | 0.23 |
| Parietal | | | |
| Left | −0.50 | −0.41 | 0.09 |
| Medial | −0.30 | −0.03 | 0.28 |
| Right | −0.30 | 0.19 | 0.47$^{*}$ |
| *Less-skilled comprehenders* | | | |
| Frontal | | | |
| Left | 0.75$^{**}$ | 0.62$^{**}$ | −0.13 |
| Medial | 0.70$^{**}$ | 1.04$^{**}$ | 0.34 |
| Right | 0.17 | 0.67$^{*}$ | 0.49$^{*}$ |
| Central | | | |
| Left | 0.41 | −0.08 | −0.49 |
| Medial | 0.44 | 0.48 | 0.041 |
| Right | −0.22 | 0.05 | 0.27 |
| Parietal | | | |
| Left | −0.06 | −0.50 | −0.44 |
| Medial | −0.02 | −0.17 | −0.15 |
| Right | −0.17 | −0.17 | 0.00 |

* $p < 0.05$.

** $p < .01$.

**** $p < 0.0001$.

Table A1

Mean frequencies, number of shared letters and second word lengths for the phonological task and mean frequencies and associative strengths of words in the semantic-word task

| Condition | Word 1 freq. | Word 2 freq. | Shared letters | Word 2 length |
|---|---|---|---|---|
| *Phonological task* | | | | |
| Homophone | 58 (92) | 63 (119) | 3.0 (0.7) | 4.4 (0.7) |
| Non-homophone | 92 (192) | 104 (237) | 2.5 (0.9) | 4.4 (0.7) |
| | Word 1 freq. | Word 2 freq. | LSA | EAT |
| *Semantic-word task* | | | | |
| Associatively related | 52 (92) | 60 (119) | 0.46 | 0.23 (0.2) |
| Categorically related | 52 (192) | 54 (237) | 0.33 | 0.008 (0.003) |
| Unrelated | 50 (87) | 69 (67) | – | – |

Note: standard deviations are shown in parentheses.

Table A3
Average amplitude differences during the 300–600 ms time window (N400) between relatedness conditions for six electrode clusters for skilled and less- skilled comprehenders for the semantic-word task

| Cluster | Associative–categorical | Associative–unrelated | Categorical–unrelated |
|---|---|---|---|
| *Skilled comprehenders* | | | |
| Central | | | |
| Left | 0.37 | 1.31** | 0.94** |
| Middle | 1.35** | 3.45**** | 2.09**** |
| Right | 0.56[a] | 2.02**** | 1.46**** |
| Parietal | | | |
| Left | 0.34 | 1.29** | 0.95* |
| Middle | 0.50 | 2.60** | 1.42** |
| Right | 0.14 | 1.89**** | 1.75**** |
| *Less-skilled comprehenders* | | | |
| Central | | | |
| Left | 0.66 | 1.48** | 0.82** |
| Middle | 0.39 | 2.48** | 2.08** |
| Right | 0.16 | 1.28** | 1.12** |
| Parietal | | | |
| Left | 0.38 | 1.54* | 1.16* |
| Middle | 0.14 | 1.90 | 1.30* |
| Right | 0.36 | 1.24* | 0.88* |

[a] $p = .06$.
* $p < .05$.
** $p < .01$.
**** $p < .0001$.

*Unrelated pairs*: aluminum–FISH, upset–ALGAE, anyone–EARTH, arabic–ALUM, bacon–AMBER, boost–ZINC, basil–CLAY, bean–MONK, beret–LARD, hassle–ICING, supple–MUSIC, boots–NORM, herpes–WOLF, booze–READ, bother–PEACE, brace–HAY, bran–TEETH, brandy–FRY, broach–AXIS, broth–BASS, calf–ARCH, cancer–BATS, cash–BLAME, cast–DEER, chamber–DICE, chaos–DILL, chess–DUSK, child–FAIL, civil–FLESH, coast–INDEX, crew–GOLD, crush–GRAVY, data–JEAN, decay–JAZZ, dew–LAVA, dough–LEAF, dust–GLOW, eagle–CLOSE, ease–COPE, east–MUD, least–PREY, horror–SELF, egg–PACK, evidence–PAD, feel–PANDA, feet–CELERY, glean–CALL, glory–SELL, leprosy–PLASTIC, lobe–MORPHINE.

*Phonological task*

*Homophone pairs*: boar–BORE, break–BRAKE, chants–CHANCE, daze–DAYS, feat–FEET, fowl–FOUL, hare–HAIR, heal–HEEL, hear–HERE, heir–AIR, jeans–GENES, maize–MAZE, meat–MEET, ore–OAR, pail–PALE, poll–POLE, rein–RAIN, sole–SOUL, steal–STEEL, suite–SWEET, tacks–TAX, tale–TAIL, wait–WEIGHT, weak–WEEK, whine–WINE, bawl–BALL, bear–BARE, bite–BYTE, board–BORED, cent–SENT, fair–FARE, flee–FLEA, hurts–HERTZ, loot–LUTE, mail–MALE, mince–MINTS, need–KNEAD, night–KNIGHT, pair–PEAR, paste–PACED, plane–PLAIN, raise–RAYS, road–RODE, role–ROLL, rose–ROWS, seas–SEIZE, site–CITE, size–SIGHS, stake–STEAK, stare–STAIR, sword–SOARED, taught–TAUT, waste–WAIST, whole–HOLE.

*Non–homophone pairs*: cove–CLIFF, mint–MOVE, bate–BARE, lips–LIST, green–GREET, seep–SEAT, melt–MEND, lobe–LONE, went–WEST, silt–SIGN, throw–THREAD, slap–SLANG, break–BREW, lake–LOCK–, harp–HAT, welt–WEAR, glint–GUILT, nest–NECK, shook–SHOOT, hills–HIDE, clown–CLOUD, east–EARN, brown–BRAIN, class–CLUB, asp–ACTS, month–MOTH, sound–SOUTH, kelp–KEY, steal–STEP, dill–DIVE, reap–REASON, crane–CREST, silt–SIZE, chance–CHAIR, ton–TOW, fowl–FOIL, oval–OBESE, toll–TOSS, rope–ROCK, sent–SELL, wilt–WEEP, beam–BEAT, dolt–DIME, ring–RIGHT, should–SCHOOL, bone–BOOK–, self–SHELF, honey–HOLY, make–MASK, moss–MODE, float–FLOUR, great–GREEN, till–TIME, lost–LOOPS.

## References

Barnea, A., & Breznitz, Z. (1998). Phonological and orthographic processing of Hebrew words: electrophysiological aspects. *Journal of Genetic Psychology, 159*(4), 492–504.

Barr, A., & Brandt, J. (1996). Word-list generation deficits in dementia. *Journal of Clinical & Experimental Neuropsychology, 18*(6), 810–822.

Beck, I. L., & Juel, C. (1992). The role of decoding in learning to read. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about reading instruction* (pp. 101–123). Newark, NJ: International Reading Association.

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.

Bird, H., Lambon Ralph, M. A., Patterson, K., & Hodges, J. R. (2000). The rise and fall of frequency and imageability: noun and verb production in semantic dementia. *Brain and Language, 73*(1), 17–49.

Cain, K., Oakhill, J., & Bryant, P. (2000). Phonological skills and comprehension failure: a test of the phonological processing deficit hypothesis. *Reading and Writing, 13*(1–2), 31–56.

Coulson, S., Federmeier, K. D., Van Petten, C. & Kutas, M. Right hemisphere sensitivity to word and sentence level context: evidence from event related brain potentials. Journal of Experimental Psychology: Learning, Memory and Cognition, 31(1), 129–147.

Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General, 132*(2), 163–201.

Federmeier, K. D., & Kutas, M. (2001). Meaning and modality: influences of context, semantic memory organization, and perceptual predictability on picture processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(1), 202–224.

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.

Hagtvet, B. E. (2003). Listening comprehension and reading comprehension in poor decoders: evidence for the importance of syntactic and semantic skills as well as phonological skills. *Reading and Writing, 16*, 505–539.

Hamm, J. P., Johnson, B. W., & Kirk, I. J. (2002). Comparison of the N300 and N400 ERPs to picture stimuli in congruent and incongruent contexts. *Clinical Neurophysiology, 113*, 1339–1350.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychological Review, 111*(3), 662–720.

Hart, L. A. (2005). A Training Study using an Artificial Orthography: Effects of Reading Experience, Lexical Quality, and Text Comprehension in L1 and L2. University of Pittsburgh dissertation.

Hillyard, S. A., & Kutas, M. (2002). Event-related potentials and magnetic fields in the human brain. In D. Charney, J. Coyle, K. Davis, & C.

Nemeroff (Eds.), *Neuropsychopharmacology: The fifth generation of progress* (pp. 427–440). Baltimore: Lippincott, Williams and Wilkins.

Kiefer, M., & Brendel, D. (2006). Attentional modulation of unconscious "automatic" processes: evidence from event-related potentials in a masked priming paradigm. *Journal of Cognitive Neuroscience, 18*(2), 184–198.

Kiss, G. R., Armstrong, C., Milroy, R., & Piper, J. (1973). An associative thesaurus of English and its computer analysis. In A. J. Aitken, R. W. Bailey, & N. Hamilton-Smith (Eds.), *The computer and literary studies*. Edinburgh: University Press.

Kramer, A. F., & Donchin, E. (1987). Brain potentials as indices of orthographic and phonological interaction during word matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*(1), 76–86.

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

Kutas, M., & Van Petten, C. K. (1994). Psycholinguistics electrified: event-related brain potential investigations. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 83–143). San Diego, CA: Academic Press.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: the latent semantic analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211–240.

Landi, N. (2006). Behavioral and electrophysiological investigations of semantic processing in skilled and less-skilled comprehenders. *Dissertation Abstracts International, 66*(09), 5116A (UMI No. 3192974).

Liu, Y., Perfetti, C. A., & Hart, L. (2003). ERP evidence for the time course of graphic, phonological, and semantic information in Chinese meaning and pronunciation decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(6), 1231–1247.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review, 109*(1), 35–54.

Manis, F. R., Seidenberg, M. S., Doi, L. M., McBride-Chang, C., & Peterson, A. (1996). On the bases of two subtypes of development dyslexia. *Cognition, 58*(2), 157–195.

Martin-Loeches, M., Hinojosa, J. A., Gomez-Jarabo, G., & Rubia, F. J. (2001). An early electrophysiological sign of semantic processing in basal extrastriate areas. *Psychophysiology, 38*(1), 114–124.

Meyler, A., & Breznitz, Z. (2005). Visual, auditory and cross-modal processing of linguistic and nonlinguistic temporal patterns among adult dyslexic readers. *Dyslexia: An International Journal of Research and Practice, 11*(2), 93–115.

Nation, K. (2005). Reading comprehension difficulties. In M. J. Snowling & C. Hulme (Eds.), *The science of reading* (pp. 248–265). Oxford: Blackwell Publishing.

Nation, K., Marshall, C. M., & Snowling, Margaret J. (2001). Phonological and semantic contributions to children's picture naming skill: evidence from children with developmental reading disorders. *Language and Cognitive Processes, 16*(2/3), 241–259.

Nation, K., & Snowling, M. J. (1998). Semantic processing and the development of word-recognition skills: evidence from children with reading comprehension difficulties. *Journal of Memory and Language, 39*, 85–101.

Nation, K., & Snowling, M. J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: evidence from semantic priming. *Cognition, 70*, B1–B13.

Nation, K., Snowling, M. J., & Clarke, P. (2005). Production of the English past tense by children with language comprehension impairments. *Journal of Child Language, 32*, 117–137.

Niznikiewicz, M., & Squires, N. K. (1996). Phonological processing and the role of strategy in silent reading: behavioral and electrophysiological evidence. *Brain and Language, 52*(2), 342–364.

Oakhill, J., & Cain, K. (2000). Children's difficulties in text comprehension: assessing causal issues. *Journal of Deaf Studies and Deaf Education, 5*(1), 51–59.

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: evidence from component skills. *Language and Cognitive Processes, 18*(4), 443–468.

Patterson, K., & Behrmann, M. (1997). Frequency and consistency effects in a pure surface dyslexic patient. *Journal of Experimental Psychology: Human Perception and Performance, 23*(4), 1217–1231.

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

Perez-Abalo, M. C., Rodriguez, R., Bobes, M. A., Gutierrez, J., & Valdes-Sosa, M. (1994). Brain potentials and the availability of semantic and phonological codes over time. *NeuroReport, 5*, 2173–2177.

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

Perfetti, C. A., & Hart, L. (2001). The lexical basis of comprehension skill. In D. 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., & Lesgold, A. M. (1979). Discourse comprehension and sources of individual differences. In P. A. Carpenter & M. A. Just (Eds.), *Cognitive processes in comprehension* (pp. 141–183). Hillsdale, NJ: Erlbaum.

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.

Plaut, D. C., McClelland, J., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychological Review, 103*(1), 56–115.

Radeau, M., Besson, M., Fonteneau, E., & Castro, S. L. (1998). Semantic, repetition and rime priming between spoken words: behavioral and electrophysiological evidence. *Biological Psychology, 48*, 183–204.

Rugg, M. D. (1984). Event-related potentials in phonological matching tasks. *Brain and Language, 23*(2), 225–240.

Sandak, R. A., Mencl, W. E., Frost, S. J., Rueckl, J. G., Katz, L., Moore, D. L., et al. (2004). The neurobiology of adaptive learning in reading: a contrast of different training conditions. *Cognitive, Affective & Behavioral Neuroscience, 4*(1), 67–88.

Segalowitz, S. J., Wagner, W. J., & Menna, R. (1992). Lateral versus frontal ERP predictors of reading skill. *Brain and Cognition, 20*(85-103).

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory, 6*(2), 174–215.

Stanovich, K. E., Siegel, L. S., & Gottardo, A. (1997). Converging evidence for phonological and surface subtypes of reading disability. *Journal of Educational Psychology, 89*(1), 114–127.

Stothard, S. E., & Hulme, C. (1995). A comparison of phonological skills in children with reading comprehension difficulties and children with decoding difficulties. *Journal of Child Psychology & Psychiatry, 36*(3), 399–408.

Thomas, M., & Karmaloff-Smith, A. (2002). Are developmental disorders like cases of adult brain damage? Implications from connectionist modeling. *Behavioral and Brain Sciences, 25*, 727–728.

Vellutino, F. R., & Scanlon, D. M. (1987). Phonological coding, phonological awareness, and reading ability: evidence from a longitudinal and experimental study. *Merrill Palmer Quarterly Special Issue: Children's reading and the development of phonological awareness, 33*(3), 321–363.

Ward, J., Stott, R., & Parkin, A. J. (2000). The role of semantics in reading and spelling: evidence for the "summation hypothesis. *Neuropsychologia, 38*, 1643–1653.

Yang, C. L., Perfetti, C. A., & Schmalhofer, F. (2005). Less skilled comprehenders' ERPs show sluggish word-to-text integration processes. *Written Language & Literacy, 8*(2), 157–181.