

Comprehension Skill and Word-to-Text Integration Processes

CHARLES PERFETTI^{1*},
CHIN-LUNG YANG¹ and FRANZ SCHMALHOFER²

¹*University of Pittsburgh, USA*
²*University of Osnabrück, Germany*

SUMMARY

We examine comprehension skill differences in the processes of word-to-text integration, the connection of the meaning of a word, as it is read, to a representation of the text. We review two 'on-line' integration studies using event related potentials (ERPs) to provide fine-grain temporal data on the word-to-text processes of adult readers. The studies demonstrate indicators for word-to-text integration and show differences in these indicators as a function of adult reading comprehension skill. For skilled comprehenders, integration processes were reflected in N400 indicators when a critical word had an explicit link to a word in the prior text and by both N400 and P300 indicators when its meaning was a paraphrase of a prior word. When forward inferences were required for subsequent word-to-text integration, effects for skilled comprehenders were not reliable. Less skilled comprehenders showed delayed and less robust ERP effects, especially when meaning paraphrase was the basis of the integration. We discuss the significance of skill differences in integration processes with a focus on the use of context-dependent word meaning as a possible source of these differences. Copyright © 2008 John Wiley & Sons, Ltd.

Text integration processes are essential to reading comprehension skill. Indeed, a widely shared view in reading education is that there are children who read words seemingly without comprehending them. Such readers are said to have good word-level and decoding skills, reading each successive word as if it were unrelated to the words already read. This description also may fit at least some adults with comprehension problems.

As a generalization on comprehension skill, however, word level skill is typically not very strong in readers who have problems in comprehension (Perfetti, 1985) and word level skill is sometimes overestimated in research that targets comprehension-specific reading problems (Perfetti, 1995). Nevertheless, the evidence seems clear that there are both children (Cain & Oakhill, 1999; Nation & Snowling, 1998; Stothard & Hulme, 1992) and adults (Hart, 2005; Landi, 2005) whose problems with comprehension are not associated with word level decoding.

Our purpose here is to examine a specific sense in which this characterization of comprehension problems might be understood: Some readers fail to effectively integrate words with prior context. Furthermore, this integration failure may involve word processing—not decoding, but the ability to link word meanings appropriately in sentence contexts.

*Correspondence to: Charles Perfetti, 644 LRDC, 3939 O'Hara Street, Pittsburgh, PA 15260, USA.
E-mail: perfetti@pitt.edu

Our examination of this issue focuses on adult readers and one particular method of studying word-by-word processes, event related potentials (ERPs). ERPs can be informative on the word-by-word reading processes of readers of different skill and can test hypotheses about word-to-text integration and other reading processes that are exposed through word-by-reading.

In the absence of word-by-word methods to study skill differences in reading, there are global and limited generalizations about reading skill. It is clear that the potential sources of comprehension problems range from basic levels of word identification (Perfetti, 1985; Perfetti & Hart, 2001) up to the higher levels of inference and co-reference processes (Long & Golding, 1993; Oakhill & Garnham, 1988) and comprehension monitoring (Baker, 1984; Garner, 1980). Within this broad range specific problems in semantic processing at the word level may also be involved, for adults (Landi, 2005) as well as for children (Nation, 2005; Nation & Snowling, 1998; for reviews, see Perfetti, Landi, & Oakhill, 2005).

Although these various components contribute to overall comprehension skill, we know less about how the processes of word-by-word reading vary with overall comprehension skill. Research that assesses only the global outcomes of comprehension processes provide mixed pictures of the causes of comprehension problems. For instance, some studies suggest that less skilled readers' failure in meaning processing might result from deficits in the quality of lexical representations (Perfetti & Hart, 2001), whereas other studies identify readers with adequate word identification abilities who show poor comprehension (Cain, 1996; Cornoldi & Oakhill, 1996). Other studies conclude that less skilled comprehenders have problems at the level of text integration in the absence of difficulties in sentence comprehension (Long, Oppy, & Seely, 1997). However, such studies do not directly illuminate the processes that occur during word-by-word reading that lead to comprehension problems.

This word-by-word reading process is what we address here. If a reader has failed to comprehend some stretch of text well, then we assume that the problem originates in local text processing. Something has gone wrong during the reading of that text, including the processes that make sense out of words. One view on what goes wrong is that the reader's useful knowledge of words is not sufficient. Accordingly, in any given text there is a risk of comprehension failing because the processes that connect word meanings to text representations do not have enough high quality word representations with which to work. The lexical quality hypothesis (Perfetti, 2007; Perfetti & Hart, 2001) proposes that at least some comprehension problems originate in low lexical quality, defined as accessible knowledge of a word's form and meaning.

We apply the lexical quality hypothesis to the integration problem as follows: As each word is read its meanings are accessed (automatically) and pruned to fit into the context, which the reader must represent as part of an understanding of the text (i.e. a mental model of the situation described by the text.). This process is recurring, with each word rapidly integrated into this continuously modified mental model. Several aspects of word knowledge are involved in this process: word form, the basis for meaning retrieval, must be stable enough to prompt retrieval of the right meaning. Word forms that are confusable on the basis of their similar spellings (e.g. *quit* and *quiet*) or identical pronunciations (*knight* and *night*) are thus potentially low quality and a reader with low quality orthographic knowledge can have problems. More problematic can be multiple word meanings that attach to a single spelling and pronunciation form (e.g. *spring*, *fair*), because the appropriate meaning is completely determined by the context.

Skilled readers have high quality representations for more words, and although they may show confusions when presented with form-similar words, they recover better than less

skilled comprehenders (Gernsbacher & Faust, 1991). In fact, when presented with the more frequent member of a pair of homophones (the highly frequent *gate* vs. the less frequent *gait*), skilled comprehenders show no confusion about its meaning, as assessed in the time to make meaning judgments (Perfetti & Hart, 2001). Less skilled comprehenders, by contrast, show such confusion in the form of interference, taking longer to decide that *gait* and *fence* are not related in meaning. This difference, we suggest, is due to the higher quality of the skilled comprehender's lexical representations, compared with those of the less skilled comprehender. That is, the skilled comprehender has more stable knowledge about both *gate* and *gait* as word forms and their connection to distinct meanings.

The lexical quality hypothesis can be understood as a general claim about word knowledge that encompasses both the stability of word forms (spelling and pronunciation) and their meanings. Form and meaning knowledge are separable, and a more specific hypothesis is that less skilled comprehenders have a semantic deficit, less accessible knowledge of word meanings (Nation & Snowling, 1998). As an explanation of comprehension problems, both this semantic deficit hypothesis and the lexical quality hypothesis are about word knowledge and its accessibility. These ideas stand in contrast to the view that poor comprehension results from defective processes, either a specific process such as making inferences or monitoring comprehension or a general deficit in working memory capacity. The evidence for a relationship between comprehension and working memory capacity is quite strong (Baddeley, Logie, & Nimmo-Smith, 1985; Crain & Shankweiler, 1988; Just & Carpenter, 1992; Perfetti & Lesgold, 1977). An active memory (as opposed to a passive storage) plays a critical role in holding information from a stretch of text (Daneman & Carpenter, 1980; Perfetti & Goldman, 1976; Seigneuric, Ehrlich, Oakhill, & Yuill, 2000). Accordingly, working memory must also be an important factor in integration processes across text segments.

Our brief review of these alternative conceptualizations of comprehension problems does not conclude that one is favoured over others on the basis of existing evidence. It is very difficult, in fact, to identify a single or even a dominant cause for comprehension problems, based on the evidence. Many different processes are involved in comprehension, and they tend to be interconnected (and interact) during reading. Any comprehension failure may be connected to multiple co-varying factors. Inferences are affected by working memory capacity, which also affects the ability to learn the meanings of words from context. But functional working memory may be affected by successful experience in reading and the growth of stable, accessible word knowledge. (See Perfetti et al., 2005, for a review.) For the present purpose, we emphasize that a critical part of reading comprehension is word-to-text integration, and that several hypotheses about causes of comprehension problems—*inference making, lexical quality, semantic deficits, working memory*—would seem to predict that word-to-text integration should be a pivotal process, one vulnerable to breakdown.

We turn now to some of the text devices that support word-to-text integration and then to a review of two recent ERP experiments that demonstrate comprehension skill differences in integration processes.

WORD-TO-TEXT INTEGRATION DEVICES

As a word in a text is identified, the reader may connect it to a continuously up-dated representation of the text. For example, consider the second occurrence of the word *spilled* in (1):

(1) The flight attendant had just served a completely full glass of red wine to the passenger when suddenly the plane hit turbulence which spilled the wine. The *spilled* wine stained the passenger's pants.

At the word *spilled* in the second sentence, the reader may immediately connect the sense of *spilled* as a modifier with the event of spilling denoted by the verb *spilled* in the preceding sentence. Alternatively, the reader can simply wait for more information, especially for the noun that a modifier such as *spilled* must modify, or even until the end of the sentence, before carrying out this integration.

The option for delaying integration is especially plausible at the beginning of a sentence. The first words of a sentence are much less constrained by prior context than are the last words of the sentence. Rather than committing to a specific, interpreted connection, the reader might encode only the word form and hold it in memory until reliable information is obtained to direct meaning integration. Nevertheless, the evidence suggests that a reader processes each word immediately, to the extent possible, rather than taking a 'wait and see' approach. The immediacy assumption, derived from studies of eye movements during reading, expresses this immediate processing idea at a general level (Just & Carpenter, 1992). At the level of syntax, very different theories about how words are attached to syntactic structures (Frazier & Clifton, 1996; MacDonald, Perlmutter, & Seidenberg, 1994) agree on the conclusion that each word is immediately attached to a syntactic phrase. Referential integration of a word's meaning with a semantic representation of the text, which maintains comprehension of the situation described by the text, may also be achieved by immediate attachment.

Although example (1) above illustrates a basic device for text integration—repeating a word already introduced in a text—the more general device for text integration is argument overlap, the occurrence of a word that is referentially bound to a previous word in the discourse. Argument overlap is the key mechanism for maintaining coherence to a minimal standard during reading, and plays a central role in models of comprehension (Kintsch, 1998). However, repeating words is not the only way to achieve referential binding, which is the more general process of making a word connect to some referent. Nor is it the most common way, because there is a constraint against repeating a word when a speaker (or writer) can assume that its referent is readily accessible. Pronouns are commonly used to link words to prior referents and have been well studied for years (e.g. Garnham, 1999; Tyler & Marslen-Wilson, 1982). We want to draw attention to two other text devices for integration, semantic paraphrase and inference making.

Example (2) illustrates semantic paraphrase, words that can refer to the same referent as some other word in a given context:

(2) The flight attendant had just served a completely full glass of red wine to the passenger when suddenly the plane hit turbulence that emptied the glass. The *spilled* wine stained the passenger's pants.

First, notice that the second sentence of (2) is the same as the second sentence of (1). What is different is that *spilled* occurs for the first time in the second sentence. It paraphrases the first sentence phrase 'emptied the glass'. Paraphrases are lexical connections between the meanings of words that produce meaning equivalence in some contexts. On this view, paraphrases are not necessarily lexical synonyms. (Note that *spill* and *empty* are not synonyms.) Although lexical synonyms are one source of paraphrase,

they do not readily appear in anaphoric relations of the kind we are discussing. This is because speakers and writers honour a general constraint to not use alternating forms co-referentially, instead choosing one and sticking with it. For example, to the extent *car* and *automobile* are synonyms, they are not used to refer to the same referent within a short stretch of discourse. (So this text seems odd: *Last week, I bought a new car. The automobile is already in the repair shop.*) The point to be made about paraphrase is that understanding words as co-referential generally requires the understanding of their context, because the context selects the meaning features of the words that allow them to be co-referential.

One could reasonably claim that some inference process is required to understand that *spilled* and *emptied* share a referent in example (2). Here, however, we want to reserve the term 'inference' to refer to more distinctive processes that add propositions to establish word-to-text integration. Example (3) below illustrates such a case:

(3) The flight attendant had just served a completely full glass of red wine to the passenger when suddenly the plane hit turbulence. The *spilled* wine stained the passenger's pants.

In (3) there are two ways an inference can lead to word-to-text integration at the occurrence of *spilled* in the second sentence. First, a forward inference can elaborate or predict during reading of the first sentence. Full glass of wine + turbulence = a spill. Thus, when *spilled* is encountered, the referent is already established through an added proposition that wine has spilled. Because it has established a referent through an added proposition, this kind of forward inference should enable word-to-text integration as easily as the explicit and paraphrasing devices illustrated above.

Suppose, however, that such a forward inference is not made during the first sentence. That sets the occasion for a backward, bridging inference to be made when *spilled* or *spilled wine* are encountered in the second sentence. This occurs when the reader notices that *spilled* or *spilled wine* does not seem to refer to anything the reader already knows about, and then makes the inference that the wine must have spilled following the turbulence. This bridging inference allows word-to-text integration and keeps the text coherent. But it appears to do so at some cost to processing at the word *spilled*. Because the inference had not been made before, it must be made as the word is read.

If we measure reading on the word *spilled*, we should detect whether there was some cost in this inference case, and in each of the other devices we have described for word-to-text integration. All costs associated with word-to-text integration at a given word can be assessed by the N400 or some other ERP component (van Berkum, Hagoort, & Brown, 1999) by comparing the three devices for word-to-text integration described above—explicit, paraphrase, and inference—with a text as in (4), which has no language by which the word *spilled* in the second sentence can be integrated with the referents introduced by the text of the first sentence.

(4) After turbulence was completely over, the flight attendant served each passenger a glass of red wine. The *spilled* wine was still lying on the floor.

Although the first sentence established some of the same referents as appear in examples (1) through (3) (turbulence, passenger, glass, red wine), its message and hence its referential world is different. There is no event of wine spilling, so the word *spilled* in sentence 2 will not be integrated with the first sentence text. Thus, (4) forms a baseline condition of difficult integration against which the success of the devices illustrated in

(1)–(3) can be assessed. In particular, we expect a large N400 measured on *spilled* in the baseline condition of (4) because integration is not possible. We expect that the explicit reinstatement of the word form *spill* (in the altered modifier form *spilled*) in example (1) will reduce the N400, because it allows a semantic fit, a word-to-text integration.

As for cases (2) and (3), we see them as making different demands on reading processes that should distinguish them in an ERP measure. The paraphrase case requires lexically based meaning processing in context. Reading that quickly activates word meanings and uses a semantic memory of the preceding text to further constrain meaning selection should show evidence of integration. However, the inference case—that is, the forward inference that would be required in the first sentence—is not a compelling one to make in context, and we think such an inference will be highly variable and probabilistic. Such an inference is sometimes made and sometimes not, according to the minimalist hypothesis, which assumes that readers generally make only those inferences necessary for text coherence and not those that merely elaborate information in the text (McKoon & Ratcliff, 1992). Of course, the occurrence of forward inferences has been the subject of much research and controversy, and there are rebuttals to the minimalist position (Graesser, Singer, & Trabasso, 1994). It is likely that texts vary in their ability to compel readers to make forward inferences. Theoretically, we see forward inferences as subject to a number of variables that make them probabilistic functions across text conditions and reader motivations, rather than the kind of text processes that are more general (Perfetti, 1993).

WORD-BY-WORD READING

The study of word-by-word reading is restricted to methods that can take measurements on the reader's response to each word. ERP methods allow word-by-word reading data without explicit responding, leading to inferences about some of the processes that occur when a word is identified and connected to its context. Shifts in EEG voltage recordings can be time-locked to stimulus events, such as the onset of a word, and related to experimental manipulations, including variations in the preceding context. In the study of meaning processes, the N400 has proved to be important. The N400 is a negative voltage shift that peaks between 300 and 500 milliseconds from the onset of a word, and is observed most strongly over central and right parietal electrodes. Its use as a marker of semantic processing rests on its sensitivity to context. When a word is not a good fit to its context (i.e. it is unexpected or semantically anomalous in that context), the N400 is observed. When the word's fit to context is good, a reduction in the amplitude of the N400 is observed; that is, the negativity is less. More generally, the magnitude of the N400 is associated with goodness of fit of a word in relation to its preceding context (Federmeier & Kutas, 2001; Kutas & Hillyard, 1980; Van Petten & Kutas, 1990).

This allows the N400 to be taken as an index of word-to-text integration difficulty. A word that is difficult to integrate with its context should produce a large N400. Any text condition that facilitates integration should reduce the N400. Indeed, the N400 has been found to be sensitive to integration in text processing (van Berkum et al., 1999). In addition to the N400, which is associated with a general semantic process, an earlier negativity, more frontal in its distribution, may be associated more specifically with tracking reference in discourse (van Berkum, Zwitserlood, Brown, & Hagoort, 2003). When the referent of a noun phrase was ambiguous (e.g. 'the girl' when the discourse had previously introduced two different girls), there is evidence for an early negativity (peak at 300–400 milliseconds)

with a frontal distribution distinct from the N400 (van Berkum et al., 2003). It remains to be seen whether the kinds of word-to-text integration we are talking about (see examples (1)–(3)) are captured by the component identified by van Berkum et al. or by the more general N400. In either case, a word-to-text integration effect is reflected in ERP shifts that arise during word reading at a point later than that associated with basic word identification (100–200 milliseconds), but within 400 milliseconds of the word's onset, relatively quickly in the context of explicit measures of word processing (e.g. naming or lexical decision).

In summary, word-to-text integration is an essential part of reading. ERPs give evidence for integration processes in (at least) the N400, which is sensitive to the degree of fit between a word and its context (or the degree of effort required to make a fit) and an additional negative component has been associated with referential processes that require additional memory resources (van Berkum et al., 1999, 2003). We can examine ERPs recorded during the reading of a critical word to compare these components under conditions that vary the text devices that support integration. Explicit word repetition should produce reduced N400s relative to a baseline condition that provides no textual support for integration. Support for integration by paraphrase, inference, and other devices can be compared with this expected effect. Finally, and critically to the present argument, is that differences in ERP markers can expose skill differences in word-to-text integration processes. We can ask whether skill differences are observed when the integration process depends on (a) explicit word repetition, (b) meaning paraphrase and (c) a forward inference.

COMPREHENSION SKILL DIFFERENCES IN ERP INDICATORS

These questions were addressed in ERP studies reported by Yang, Perfetti and Schmalhofer (2005, 2007) using materials that correspond to examples (1)–(4). Table 1 displays the four conditions of the experiment with new examples. To review, if a word has occurred in the previous text, even if not in the exact form, its integration should readily occur. This defines the *explicit* reference condition illustrated in (1). If a word has not occurred in the preceding text, but in that context is related to a word that did occur, there will be a reduction in the N400, only if a lexico-semantic connection is made. This defines the *paraphrase* condition illustrated in example (2). If a word has not occurred in the preceding text, but could have been implicitly retrieved as part of an inference, there will be a reduction in the N400 only if the inference has been made. This defines the *inference* condition of example (3). The N400 will be reduced only by a forward inference, one made on the basis of the preceding sentence. Finally, when the word has no plausible antecedent in the first sentence, its integration is not possible. This defines the *baseline* condition of example (4), and should produce the largest N400.

We expected the explicit condition to provide support for integration, because the repetition of a word, even in a variant form, invites a referential connection. So, in Table 1, the word *hospital* in the second sentence is readily connected to a referent already introduced in the first sentence—the hospital. (It is possible that a N400 reduction here could reflect a lexical repetition effect rather than a semantic-referential effect.) Furthermore, on the assumption that word-based processing dominates reading, we reasoned that paraphrase would support integration. An encounter with *hospital* in sentence (2) would prompt a link to the referent implied by *emergency room* in sentence (1). In the

Table 1. Experimental materials illustrated for the word *hospital*

Condition	Sentence context
Explicit	Allen's baby became violently ill, so Allen got the baby in the car and rushed off to the hospital. The hospital had a long waiting line
Paraphrase	Allen's baby became violently ill, so Allen got the baby in the car and rushed off to the emergency room. The hospital had a long waiting line
Inference	Allen's baby became violently ill, so Allen got the baby in the car and rushed off. The hospital had a long waiting line
Baseline	Allen rushed off to work, when his wife was no longer feeling very ill. The hospital that she finally went to was very crowded

case of inferences, we reasoned that making predictive inferences was not a routine part of reading, and may not lead to an N400 reduction for all readers.

ERP methods

In ERP studies of reading, the text must be presented one word at a time, allowing ERPs to be time-locked to the onset of a specific word without interference from head and eye movements. In the examples of Table 1, we observe the ERPs from the onset of the critical word *hospital*.

In the first study (Yang et al., 2007), the focus was on the basic questions of text integration and whether ERPs would provide evidence that integration occurred immediately even across a sentence boundary. Accordingly, the participants were a group of skilled comprehenders from a university population, mean age 18.5. All participants had scored high on the Nelson–Denny test of reading comprehension ($M = 89\%$, $SD = 7\%$).

In the experiment, participants read the 120 two-sentence passages of the sort shown in Table 1, with each word displayed for 600 milliseconds. The first content word of the second sentence was always the critical word for measurement; that is, although EEGs were continuously recorded during reading, the key experimental measures were the ERP recordings from the onset of the critical word. This word was the same across all four conditions (explicit, paraphrase, inference and baseline), allowing a comparison of these conditions. These critical words were often nouns, as in the *hospital* example, but some were modifiers, as in the sentences with *spilled* illustrated in the first section of this paper. The recordings were made by a 128 channel electrical geodesics system (Tucker, 1993) and were referenced to the average-reference vertex (Lehmann & Skrandies, 1980). Additional details are in Yang et al. (2007).

Skilled comprehenders

Figure 1 shows the grand mean amplitude recorded at the Cz (vertex) electrode for each of the four text conditions for skilled comprehenders (top) and less skilled comprehenders (bottom). The N400 can be seen as a negative (downward) shift that extends from about 350 to 550 milliseconds within the typical N400 range.

Consider first the skilled comprehenders. The baseline condition produced the largest negativity in this time window, as expected. The explicit and paraphrase conditions produced significantly less negativity. The inference condition was not significantly

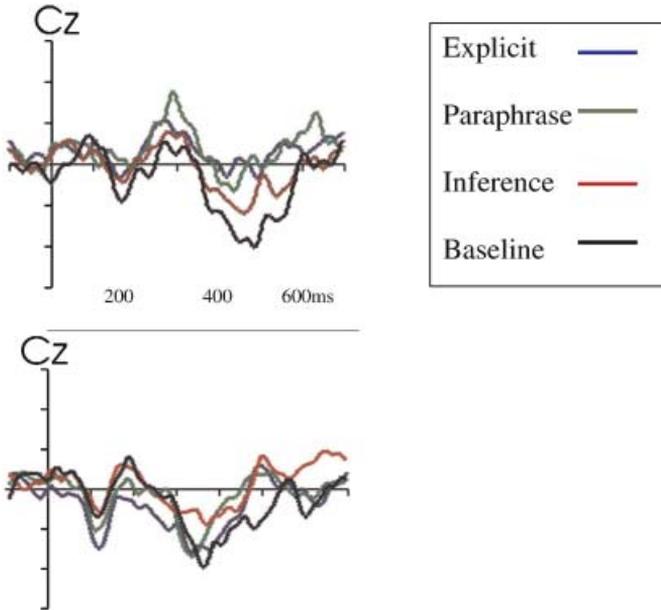


Figure 1. Mean amplitudes recorded at electrode Cz for from the onset of the critical word for four text conditions. The upper figure is for skilled comprehenders. The lower figure is for less skilled comprehenders. Positive is up, negative down. Data are adapted from Yang et al. (2007) and Yang et al. (2005), respectively

different from the baseline, although one can see that its negativity appears intermediate with respect to the other conditions. This pattern of significant N400 reductions in the explicit and paraphrase conditions, but not the inference condition, is evidence of word-to-text integration that was replicated in statistical tests based on larger samples of the 128 electrodes. (The effects were most apparent over central and parietal sites, typical for the N400.) This pattern was also revealed by a principle components analysis (PCA) that extracts factors from the raw data across the entire recording time window. Other ERP shifts occurred at the critical word as well, and the PCA identified three factors starting at 110 milliseconds (with PCA peaks at 162, 314 and 516 milliseconds) after onset of the word that were influenced by experimental conditions. The earliest factor related to word processing rose from 110 milliseconds with a peak PCA loading at 162 milliseconds. This PCA factor corresponded to an ERP component visible as a posterior negativity for the baseline condition (peak around 160 milliseconds, an N200) with reduced (marginally significant) negativity for explicit and paraphrase conditions.

An especially interesting additional effect is visible in the recording for skilled comprehenders in Figure 1. The paraphrase condition produced a significant positivity at around 300 milliseconds (a P300). This positivity was general across anterior and central electrodes. For text reading, this positive shift may be linked to the interpretation of the P300 observed in other tasks, specifically as an indicator of memory updating in situations involving novel stimuli (Donchin & Coles, 1988a, 1988b). We may generalize memory updating to capture the idea that a cognitive model of the environment is modified and updated as a function of incoming information (Donchin & Coles, 1988a). In the case of word-to-text integration, this updating process is responding to a novel word (a paraphrase)

that can be associated conceptually with the preceding text representation (Keenan & Jennings, 1995). In effect, encountering *hospital* at the beginning of the second sentence 'updates' the referent established by the phrase *emergency room* in the first sentence.

Here, then, is what the results seem to say about word-to-text integration for skilled comprehenders. ERPs suggest that skilled comprehenders immediately integrate a word into its context when this is possible, even across a sentence boundary. ERP shifts show word processing effects that may be associated with text conditions by 200 milliseconds, in a time window and a distribution that is associated with word identification. However, the clearest separation by conditions occurs somewhat later, with a separation of the paraphrase condition from all others at 300 milliseconds (P300) and a separation of explicit and paraphrase conditions from the others at around 500 milliseconds (N400). Especially interesting is that when a critical word is a paraphrase of a possible antecedent evidence for integration is obtained not only in an N400, but also in a P300. We think the earlier indicator may reflect a lexical linkage process through memory. The N400 indicator then reflects the referential matching or congruence process. In the absence of adequate antecedents for a word (the baseline and inference conditions) no easy referent matching is possible.

Less-skilled comprehenders

The study of less skilled comprehenders used the same 120 passages and procedures used in the study of skilled comprehenders. The participants were from the same university population, but were chosen to be below our sample norm on the Nelson–Denny reading comprehension test ($M = 70\%$, $SD = 9\%$). (The distribution of comprehension scores did not overlap those of the skilled group.) These were not poor readers in an absolute sense, merely readers below the mean comprehension score for this population.

Mean amplitudes for the vertex electrode (Cz) are shown in the bottom part of Figure 1. The main conclusion from the ERP data is that, compared with the skilled comprehenders, the less skilled comprehenders showed a pattern of results that suggested 'sluggish' word-to-text integration processes (Yang et al., 2005). For example, little or no reduction for the explicit and paraphrase conditions is seen in the N400 time window for the vertex electrode (Figure 1), instead being delayed and smaller. The inference condition produced the only reduction during the typical N400 time window.

To be clear, the ERP results for less skilled comprehenders showed a general similarity to that of the skilled comprehenders. For example, the PCA identified three factors across the reading time on the critical word, just as it did for the skilled comprehenders. However, the temporal peaks identified with these three factors were delayed compared with the peaks for the skilled comprehenders: 224, 360 and 656 milliseconds (compared with 162, 314 and 516 milliseconds, respectively for skilled comprehenders). These temporal differences, from the early time windows that reflect word identification through the later windows that reflect integration, suggest a generally slower word processing for less skilled comprehenders. The less skilled comprehenders showed no ERP effects related to experimental conditions prior to 300 milliseconds.

On the specific question of whether the less skilled comprehenders show an N400 effect, the answer is yes, but with interesting differences. The baseline condition produced a pronounced negativity from about 300 milliseconds that peaked at 400 milliseconds on the central vertex (see Figure 1) and central parietal electrode and somewhat later on right parietal and temporal electrodes. This indicates a fairly typical N400 effect. However, the

key question concerns the reduction of this N400 when the text conditions support integration. Here, the pattern departs from that of skilled comprehenders. The reduction for explicit and paraphrase conditions did not occur in the central Cz electrode as it did for skilled comprehenders (Figure 1). Instead, the N400 reduction as a function of text integration occurred mainly at the right posterior sites and only for the explicit condition. This we interpret as a simple effect of word repetition, because all other effects occurred later, peaking at 656 milliseconds, when both explicit and inference conditions separated from the baseline condition. The paraphrase condition showed only a late and modest (only at the left anterior site) separation from baseline.

The pattern suggests two phases, one of word identification, which is affected only by word repetition, and a later phase of word-to-text integration that follows word identification. In contrast to the results for skilled comprehenders, there was no suggestion of an early separation of conditions at the time associated with word identification. Two other departures from the results for skilled comprehenders are noteworthy. First, the less skilled comprehenders, unlike the skilled comprehenders, did show a late inference effect relative to baseline. Second, and most interesting, they showed the paraphrase effect only weakly and somewhat later than the skilled comprehenders. The first result argues against the hypothesis that, in this college student population, less skilled comprehenders have a problem making forward inferences. The second result suggests that using word meanings in context *is* a problem for this group. Overall the ERP pattern suggests sluggish word processing during text reading for less skilled comprehenders. The processes that identify words and link their meanings to previously read text are slowed and show less effective use of meaning relations among words.

To illustrate the contrast between skilled and less skilled comprehenders in using paraphrases for integration, Figure 2 presents topographical maps. These maps show time slices during the reading of a word in the paraphrase condition—for example, reading the word *hospital* when it has been preceded by the phrase *emergency room*. The voltage shifts are revealed over time, with positivities as light and negativities as dark. The course of integration is signalled by positivities in central and parietal areas especially.

The skill contrast is seen in the rapid and robust positivities for skilled comprehenders compared with the negativities for less skilled comprehenders that give way slowly to positivities. Thus, integration that depends on context-dependent meaning similarity between words produced only sluggish integration for the less skilled comprehenders.

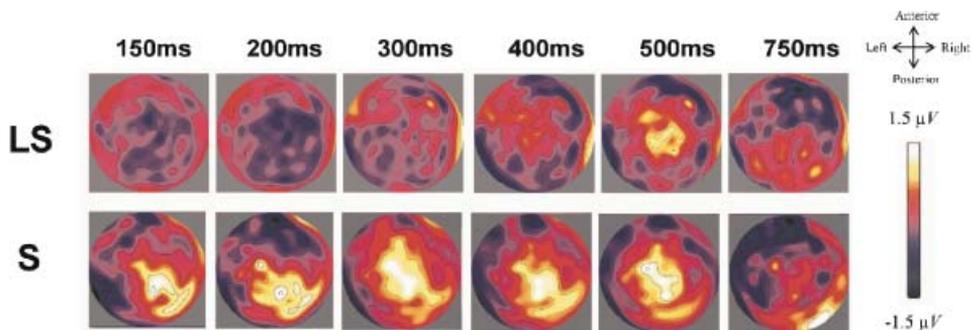


Figure 2. Topographical maps showing growth of positivities (light colour) in paraphrase condition over the course of word processing from 150 to 750 milliseconds. Less skilled comprehenders (LS), top panel, show little growth in positivities compared with skilled comprehenders (S), bottom panel

Yang et al. (2005) present topographic maps for other conditions, which show a similar skill difference for the explicit condition.

DISCUSSION

The results of these studies join a number of others that demonstrate the value of ERPs as indicators of word-to-text integration. In addition to the N400, which has become the standard indicator for context-related meaning processes generally, additional indicators are emerging. One is the 300 milliseconds frontal negativity identified by van Berkum et al. (1999) and another may be the medial-central 300 milliseconds positivity for paraphrases found by Yang et al. (2007), especially for skilled comprehenders. Both may be indicators of more specific word-to-text integration processes. The first may indicate searches for referents and the second may indicate memory operations that connect a word to a related word in memory. The interpretation that paraphrases work through lexically stimulated memory retrieval is strengthened by fMRI results described in Schmalhofer and Perfetti (2007).

For generalizations to other reading situations, one might wonder whether the word-by-word display required by ERPs produces results not generalizable to normal reading. More work is needed, especially research that compares eye movement and ERP measures on comparable texts, to address this. However, it is important to emphasize the value that ERP measures bring to the study of comprehension by allowing multiple processing events to be inferred from a single measurement taken on a word. Such an inference is not possible from behavioural studies, including eye-movement studies. In our case, results from the paraphrase condition suggested two overlapping processes indicated by the P300 and N400. The first was a memory retrieval stimulated by reading the paraphrase word, and the second was a referential fit made from this word's meaning to the text representation. This two-phase hypothesis may prove to be incorrect with further work, but the point is the potential for observing multiple effects over a brief time period. This potential also is seen in the separation of an early component, prior to 200 milliseconds, that marks basic word identification and the beginning of a semantic process that leads to integration.

Interestingly, multiple indicators are also seen in the difference between skilled and less skilled comprehenders. The less skilled group produced two indicators when they read a word that appeared in the previous sentence (the explicit condition). An early indicator seemed to reflect only the repetition of the word with a later indicator reflecting its integration into context. Skilled comprehenders showed only a single indicator of repetition, which we inferred signalled word-to-text integration.

Turning now to general text comprehension issues, the ERP studies draw attention to the role of word reading. Word-to-text integration is largely a lexico-semantic process that begins with the identification of a word from its orthographic input and, in overlapping phases, moves to the activation of associated meanings, both from the reader's lexicon and from the temporary memories that store words and referential situations from the text. The processes are very rapid and highly overlapping in skilled reading, as evidenced by the early separation of conditions in the experiment with skilled comprehenders. At a time window associated with basic word identification, we see differences that depend on what has come previously in the text.

Outside the scope of this general picture is the case of forward inferences, for which a lexically based connection is not present. It appears that, if skilled comprehenders make

forward inference in sentences like those we studied such inferences do not necessarily create a referent to which a word can be connected. It is possible the forward inference is not made at all in a reliable way. However, we do not conclude that readers do not make forward inferences. We conclude only that they do not make them routinely and consistently. We do think that inferences that are implicit in a word's meaning might behave differently (enabling inferences that allow word-to-text integration) and we are examining these in other experiments.

An important question that merits more research is the nature of successful paraphrase; that is, word variation in text that allows easy referential connections. Successful paraphrases are clearly not about synonyms. What we have termed paraphrase represented a wide range of semantic relations in these studies. The example of *emergency room* being an antecedent for *hospital* is actually part-whole semantic relationship. Emergency rooms are in hospitals, so the mention of an emergency room enables immediate use of the word *hospital*. But other relations were present as well, as illustrated by this example: *Brad fumbled through the dark until he located the box of matches and struck one. After lighting the match, it was easier to see.* If one strikes a match successfully, he or she has lit the match. Still, *strike* and *light* are not synonyms. They can refer to the same event in context. For a final example: *The male robin climbed out of the nest, spread his wings, and took off. He flew south where it was warmer this time of year.* Again, *fly* and *take off* are not synonyms, if synonyms are defined as words that can be interchanged in most contexts. Rather the proposition that an able-to-fly object took off gives license to the use of *fly*. (The object is a bird here, but airplane would work as well.)

In general, the meaning processes we have studied under the label of 'paraphrase' require the selection of word meaning appropriate for context. Thus we may be tapping not only the importance of word knowledge but also the ability to use context in meaning processing. One might prefer to think of our paraphrase processes as a kind of inference made across word meanings that are interpreted in reference to a single situation. If so, such inferences are very different from the kind of elaborative predictive inference for which we obtained no strong evidence for skilled comprehenders.

Although the phrase 'integration process' may suggest an active strategy on the part of the reader, it is possible that word-to text integration instead may be a passive process, at least in part. There is no reason to assume that readers in this study were actively searching for word-to-text links as they read a word. Instead, consistent with resonance models of memory-based processes in reading (O'Brien, Rizzella, Albrecht, & Halleran, 1998), the key integration processes may occur when reading a given word activates memory traces laid down by prior reading. To the extent that skilled readers have knowledge of the words they read and the concepts they encode and to the extent that they have adequate working memory resources, immediate word-to-text integration can occur readily based on the meanings of words.

Finally, to return to the characterization of comprehension skill, we emphasize that global differences in comprehension skill, as assessed by a standardized comprehension test, may arise in local word-by-word processing. To explain the results of the studies reviewed here, one needs only to refer to local text processes extending just beyond a single sentence; no higher level strategic processes need to be invoked. Nevertheless, the differences in ERP patterns were substantial and dramatically so in the temporal dynamics of the topographic maps. Word-by-word processes vary with reader skill and this variance has consequences for comprehension. We do not assume a single ultimate cause in this account of word-to-text integration as part of comprehension skill. Other situations using

different text and strategy manipulations might expose differences in higher-level comprehension strategies.

Nevertheless, the studies reviewed here show that small differences in the temporal dynamics of word-to-text integration are associated word-by-word with reading skill. This suggests that the causes of low comprehension skill in this range of skill include local text processes at the word level. Elaborating further the causes of skill-related word-to-text integration differences—working memory and word knowledge are arguably involved—requires research that can separate the causal factors more thoroughly. Working memory associations with adult reading skill are well established (Daneman & Carpenter, 1980) and extend to the ability to use context to understand word meanings (Daneman & Green, 1986). Especially relevant for our ERP studies is the finding that individual differences in working memory capacity are associated with the ERP indicators of the use of sentence context (Van Petten, Weckerly, McIsaac, & Kutas, 1997). The ability to retain in memory short stretches of text is important for being able to link a new word to this memory.

Our hypothesis is that, aside from working memory differences, readers vary in the flexibility of their semantic representations of words, a dimension of lexical quality (Perfetti & Hart, 2001), and that this variability has consequences for retrieval of word meanings and the selection of context-appropriate meanings. It is possible that such differences in word knowledge derive from differences in working memory as argued by Gathercole and Baddeley (1993). Sorting out primary causal relations in word-level skill differences is difficult and remains to be done, including whether differences in word knowledge result mainly from working memory differences, from variations in language experiences, or from something else. However, the ultimate source of differences in word knowledge is not the only question; indeed, not the main question for understanding skill in word-by-word comprehension. For this we need to better understand the nature and variability of readers' word knowledge, their memory representations of short stretches of text, and how the two mutually interact during reading.

REFERENCES

- Baddeley, A., Logie, R., & Nimmo-Smith, I. (1985). Components of fluent reading. *Journal of Memory and Language*, 24, 119–131.
- Baker, L. (1984). Spontaneous versus instructed use of multiple standards for evaluating comprehension: Effects of age, reading proficiency, and type of standard. *Journal of Experimental Child Psychology*, 38, 289–311.
- Cain, K. (1996). Story knowledge and comprehension skill. In C. Cornoldi, & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and interventions* (pp. 167–192). Mahwah, NJ: Erlbaum.
- Cain, K., & Oakhill, J. V. (1999). Inference ability and its relation to comprehension failure in young children. *Reading and Writing*, 11, 489–503.
- Cornoldi, C., & Oakhill, J. (Eds.), (1996). *Reading comprehension difficulties: Processes and interventions*. Mahwah, NJ: Erlbaum.
- Crain, K., & Shankweiler, D. (1988). Syntactic complexity and reading acquisition. In A. Davison, & G. M. Green (Eds.), *Linguistic complexity and text comprehension: Readability issues reconsidered* (pp. 167–192). Hillsdale, NJ: Erlbaum.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–466.
- Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–466.

- Donchin, E., & Coles, M. G. H. (1988a). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*, 355–372.
- Donchin, E., & Coles, M. G. H. (1988b). On the conceptual foundation of cognitive psychophysiology. *Behavioral and Brain Sciences*, *11*, 406–417.
- 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*, 202–224.
- Frazier, L., & Clifton, C., Jr. (1996). *Construal*. Cambridge, MA: MIT Press.
- Garner, R. (1980). Monitoring of understanding: An investigation of good and poor readers' awareness of induced miscomprehension of text. *Journal of Reading Behavior*, *12*, 5–63.
- Garnham, A. (1999). Reference and anaphora. In S. Garrod, & M. J. Pickering (Eds.), *Language processing* (pp. 335–362). East Sussex, UK: Psychology Press.
- Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. New York: Psychology Press.
- Gernsbacher, M. A., & Faust, M. E. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 245–262.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*, 371–395.
- Hart, L. (2005). A training study using an artificial orthography: Effects of reading experience, lexical quality, and text comprehension in L1 and L2. Unpublished doctoral dissertation, University of Pittsburgh.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122–149.
- Keenan, J. M., & Jennings, T. M. (1995). Priming of inference concepts in the constructionist integration model. In C. A. Weaver, S. Mannes, & C. R. Fletcher (Eds.), *Discourse comprehension: Essays in honor of Walter Kintsch* (pp. 233–244). Hillsdale, NJ: Erlbaum.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, UK: Cambridge University Press.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203–205.
- Landi, N. (2005). Behavioral and electrophysiological investigations of semantic processing in skilled and less-skilled comprehenders. *Dissertation Abstracts International*, *66*, 5116A. (UMI No. 3192974).
- Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and Clinical Neurophysiology*, *48*, 609–621.
- Long, D. L., & Golding, J. M. (1993). Superordinate goal inferences: Are they automatically generated during comprehension? *Discourse Processes*, *16*, 55–73.
- Long, D. L., Oppy, B. J., & Seely, M. R. (1997). Individual differences in reader's sentence- and text-level representations. *Journal of Memory and Language*, *36*, 129–145.
- MacDonald, M. C., Perlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676–703.
- McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, *99*, 440–446.
- Nation, K. (2005). Reading comprehension difficulties. In M. J. Snowling, & C. Hulme (Eds.), *The science of reading* (pp. 248–265). Oxford: Blackwell.
- 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.
- Oakhill, J., & Garnham, A. (1988). *Becoming a skilled reader*. New York: Blackwell.
- O'Brien, E. J., Rizzella, M. L., Albrecht, J. E., & Halleran, J. A. (1998). Updating situational models. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *24*, 1200–1210.
- Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*, 357–383.
- Perfetti, C. A. (1985). *Reading ability*. New York: Oxford Press.
- Perfetti, C. A. (1993). Why inferences might be restricted. *Discourse Processes*, *16*, 181–192.

- Perfetti, C. A. (1995). Cognitive research can inform reading education. *Journal of Research in Reading, 18*, 106–115.
- Perfetti, C. A., & Goldman, S. R. (1976). Discourse memory and reading comprehension skill. *Journal of Verbal Learning and Verbal Behavior, 14*, 33–42.
- Perfetti, C. A., & Hart, L. (2001). The lexical bases of comprehension skill. In D. Gorfien (Ed.), *On the consequences of meaning selection* (pp. 67–86). Washington, DC: American Psychological Association.
- Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling, & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227–247). Oxford: Blackwell.
- Perfetti, C. A., & Lesgold, A. M. (1977). Discourse comprehension and sources of individual differences. In M. A. Just, & P. A. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 141–183). Hillsdale, NJ: Erlbaum.
- Schmalhofer, F., & Perfetti, C. A. (Eds.), (2007). *Higher level language processes in the brain: Inference and comprehension processes*. Mahwah, NJ: Erlbaum.
- Seigneuric, A. S., Ehrlich, M. F., Oakhill, J. V., & Yuill, N. M. (2000). Working memory resources and children's reading comprehension. *Reading and Writing, 13*, 81–103.
- Stothard, S. E., & Hulme, C. (1992). Reading comprehension difficulties in children: The role of language comprehension and working memory skills. *Reading and Writing, 4*, 245–256.
- Tucker, D. M. (1993). Spatial sampling of head electrical fields: The geodesic sensor net. *Electroencephalography and Clinical Neurophysiology, 87*, 154–163.
- Tyler, L., & Marslen-Wilson, W. (1982). Processing utterances in discourse context: Online resolution of anaphors. *Journal of Semantics, 1*, 297–314.
- van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience, 11*, 657–671.
- van Berkum, J. J. A., Zwitserlood, P., Brown, C. M., & Hagoort, P. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research, 17*, 701–718.
- Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition, 18*, 380–393.
- Van Petten, C., Weckerly, J., McIsaac, H. K., & Kutas, M. (1997). Working memory capacity dissociates lexical and sentential context effects. *Psychological Science, 8*, 238–242.
- Yang, C. -L., Perfetti, C. A., & Schmalhofer, F. (2007). Event-related potential indicators of text integration across sentence boundaries. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 55–89.
- Yang, C. L., Perfetti, C. A., & Schmalhofer, F. (2005). Less skilled comprehenders' ERPs show sluggish word-to-text integration processes. *Written Language & Literacy, 8*, 233–257.