Less skilled comprehenders' ERPs show sluggish word-to-text integration processes

Chin Lung Yang[1], Charles A. Perfetti[1] and
Franz Schmalhofer[2]
[1]Learning Research and Development Center, University of Pittsburgh /
[2]Institute of Cognitive Science, University of Osnabrück

We examined the word-to-text integration processes of less skilled comprehenders using ERPs recorded during text reading. The first sentence of each text controlled the accessibility of an antecedent referent for a critical word, which was the first content word of the second sentence. In the explicit condition, the critical word had occurred in the first sentence; in the paraphrase condition, a word or phrase similar in meaning had occurred in the first sentence; in the inference condition, a referent could have been established during the first sentence only if the reader made a forward inference; a baseline condition provided no obvious antecedent for the critical word. PCA, topographic results, and mean amplitude analyses converged on a picture of integration difficulty. Integration effects emerged in the expected mid-latency ranges for the explicit and inference conditions. The pattern of effects differed from that of skilled comprehenders, who, in another study, showed earlier integration effects for explicit and paraphrase conditions, but not reliably for the inference condition. Paraphrase effects were especially weak and late occurring for less skilled comprehenders. Compared with skilled comprehenders, less skilled comprehenders show slow word-to-text integration processes.

Difficulties in comprehension can arise from any one or more of the component processes that must be integrated to produce skilled comprehension (Perfetti, 1999). These sources of comprehension problems range from basic levels of word identification (Perfetti, 1985; Perfetti & Hart, 2001) up to the higher levels of inference processes (Long & Golding, 1993; Oakhill & Garnham, 1988) and comprehension monitoring (Baker, 1982; 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 children (Nation & Snowling, 1998; for reviews, see Perfetti, Landi & Oakhill, 2005; Nation, 2005).

Although there is evidence that all these 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 provides mixed pictures of the causes of comprehension problems. Whereas some studies suggest that less skilled readers' failure in meaning processing might result from deficits in the quality of lexical representations (Perfetti & Hart, 2001), other studies identify readers with adequate word identification abilities who show poor comprehension (Cain, 1996; Cornoldi & Oakhill, 1996). Still 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.

The study we report here is aimed to do just that, to expose problems that occur in word-by-word reading that might contribute to reduced skill in comprehension. Our study examines reader's integration of a word meaning with a preceding text segment by using event related potentials (ERPs) recorded during reading. ERP recordings taken on a given word in a text can provide information on how easily the reader is able to integrate the word with what has appeared in the text to that point. We adopted the stimuli, design, and procedures of a recent ERP study by Yang, Perfetti & Schmalhofer (under review). This study tested the immediacy hypothesis, which is that readers attempt to integrate each word into a representation of the text. We interpret this process as one of word-to-referent mapping, insofar as we assume that the reader tries to assign each word to some referent that might have been established in a model of the situation (Kintsch & van Dijk, 1978) described by the text.

Yang et al. (under review) found that even at the beginning of a sentence, skilled readers showed evidence for this kind of integration despite the fact that the only relevant context came from the preceding sentence. Participants read coherent two-sentence passages (see Table 1) that varied the accessibility of an antecedent referent for the first content word of the second sentence (the critical word): In an *explicit* condition, a word from the first sentence was repeated (with morphological variation); in a *paraphrase* condition, the first sentence contained a word or phrase that was semantically similar to the critical word; and in an *inference* condition, the situation described by the first sentence could trigger an inference that established an antecedent for the critical word. In a baseline condition, the first sentence contained no apparent antecedent for the critical word.

Table 1. Sample passages in each experimental condition. The critical word (explosion), whose reference has been established to differential degrees of availability in the context representation is in underlined bold at the beginning of the second sentences. These are the critical words where the ERPs were extracted from their onset.

INTEGRATION TYPE	Sample Passage 1
Explicit	After being dropped from the plane, the bomb hit the ground and exploded. The **<u>explosion</u>** was quickly reported to the commander.
Paraphrase	After being dropped from the plane, the bomb hit the ground and blew up. The **<u>explosion</u>** was quickly reported to the commander.
Inference	After being dropped from the plane, the bomb hit the ground. The **<u>explosion</u>** was quickly reported to the commander.
Baseline	Once the bomb was stored safely on the ground, the plane dropped off its passengers and left. The **<u>explosion</u>** was quickly reported to the commander.
	Sample Passage 2
Explicit	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 **<u>spilled</u>** wine stained the passenger's pants.
Paraphrase	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 **<u>spilled</u>** wine stained the passenger's pants.
Inference	The flight attendant had just served a completely full glass of red wine to the passenger when suddenly the plane hit turbulence. The **<u>spilled</u>** wine stained the passenger's pants.
Baseline	After turbulence was completely over, the flight attendant served each passenger a glass of red wine. The **<u>spilled</u>** food was still lying on the floor.

The ERP results on the critical word showed evidence for immediate integration, within the first ~300 ms from the onset of the critical word. The pattern of results suggested that lexico-semantic (explicit condition) and conceptual processes (paraphrase condition) facilitate word-to-text integration. For example, an N400 effect observed for the baseline condition was reduced in the explicit and paraphrase conditions, but not reliably in the inference condition. The conclusions were that, for skilled comprehenders, word-to-text integration is facilitated by lexical and semantic overlap across sentence boundaries and that forward inferences are not robust enough to reliably facilitate integration.

Whereas Yang et al. (under review) found immediate integration effects for skilled comprehenders, it is possible that less skilled readers show a less effec-

tive word-to-text integration process. According to Perfetti and Hart's (2001) lexical quality hypothesis, a critical source of comprehension skill is the quality of a reader's word knowledge, which allows context-appropriate retrieval and selection of word meanings. Less skilled comprehenders, on this hypothesis, may show some difficulties in the immediate integration of a word's meaning with a preceding context. Furthermore, integration opportunities arise later in a sentence, allowing the possibility of integration delay (Schmalhofer, & Perfetti, in press). Thus, the present study, in combination with the results of Yang, et al. (under review), provides a comparison of the word-to-text integration processes of more and less skilled comprehenders.

ERPs provide fine grain timing data without requiring an explicit response from the reader. Especially useful for the study of semantic processing is the N400 effect, a negative voltage shift between 300 and 500 milliseconds after the onset of a word. The less a word fits into the semantically established context, the larger the N400 amplitude (Federmeier & Kutas, 2001; Kutas & Hillyard, 1980; Van Petten & Kutas, 1990). Furthermore, the N400 has been found to be sensitive to integration in text processing (van Berkum, Brown, Zwitserlood, Kooijman & Hagoort, 2005; van Berkum, Zwitserlood, Brown & Hagoort, 2003; van Berkum, Hagoort & Brown, 1999), including inferences (St. George, Mannes & Hoffman, 1997). Additional ERP shifts in frontal electrodes have been shown to relate to memory processing related to referent integration (Kluender & Kutas, 1993; King & Kutas, 1995), and to cognitive tasks that demand recruitment of additional resources for mental operations (Karis, Fabiani & Donchin, 1984; Rosler, Heil & Hennighausen, 1995; Ruchkin, Johnson, Mahaffey & Sutton, 1988). Thus, we examine the posterior N400 for basic evidence of word-to-text integration and additional anterior ERPs for evidence of memory processes that can support more effortful integration.

Several previous ERP studies have made comparisons of groups differing in reading skill or working memory capacity. Kutas and King (1996), for example, found that low skill comprehenders showed a greater left-lateralized positivity over occipital sites compared with skilled comprehenders, suggesting a difference in the resources required for word level processing. Other ERP differences between skill groups have been reported for difficult syntactic processes (King & Kutas, 1995; Vos & Friederici, 2003) and sensitivity to text anomalies (Van Petten, Weckerly, McIsaac & Kutas, 1997). More closely related to our study is an inference study by St. George et al. (1997). They reported that low span participants showed N400 reductions for bridging inferences whereas high span participants showed N400 reductions for elaborative as well as bridging inferences.

The study we report here tests less skilled comprehenders' word-to-text integration based on the extent to which a first sentence provides an antecedent for the first content word of a following sentence. We compare these results with those on the identical texts for skilled comprehenders from Yang et al. (under review).

Experiment

We used the coherent two-sentence passages from Yang et al. (under review) as shown in Table 1. The accessibility of an antecedent referent for the first content word of the second sentence varied across four conditions. Two conditions were based on lexico-semantic connections across the sentence boundary, explicit and paraphrase. For integration to occur easily in the inference condition, the reader would need to make a situationally-based forward inference during the reading of the first sentence in order to establish an antecedent. Alternatively, in the absence of the predictive inferencing, integration would occur by a bridging inference from the second sentence back to the first (Singer & Halldorson, 1996). In a baseline condition, the critical word in the second sentence had no explicit nor implicit antecedent in the first sentence, and thus could not be easily integrated with it. Note, however, that there is no anomaly in this baseline condition and that various continuations after the first content word are meaningful and plausible in the context.

ERPs were measured on the first content word of the second sentence, which was identical across all conditions. The baseline condition was expected to produce an N400 because there is no basis for the immediate integration of the critical word with the preceding context. To the extent that any other condition facilitates integration relative to the baseline, a reduction of the N400 should occur, as it did for the skilled readers in the explicit and paraphrase conditions of Yang et al. (under review).

Yang, et al. (under review) also found that the paraphrase condition produced an earlier positivity (P300) that they interpreted as an active memory retrieval induced by semantic-conceptual processing. In contrast, the reduced N400 of the explicit condition could be interpreted as reflecting a passive memory resonance, which would be sufficient for linking a repeated word with its antecedent. Thus, we might expect that less skilled comprehenders will show a reduced N400 for the explicit condition based on explicit lexical matching, but not for the paraphrase condition if their lexico-semantic processing skills are weak. As to the inference condition, we would expect a reduced N400 to

the extent that readers drew a forward inference in the first sentence — that the bomb exploded — establishing an antecedent for *explosion* in the second sentence, a result not obtained for skilled comprehenders by Yang et al. (under review). Finally, to the extent that additional processing resources are recruited to support integration, ERP effects could be observed in the anterior brain regions.

Methods and Materials

Participants

Eighteen undergraduate students of the University of Pittsburgh (9 females) between the ages of 18 and 56 (mean, 22.6; SD, 9.3) participated in the study. We recruited subjects from a pool that had previously been identified as less skilled comprehenders using the Nelson-Denny comprehension test (mean N-D accuracy percent score, 0.70; SD, 0.09). Participants were right-handed native English speakers with normal or corrected-to-normal vision. Each participant received a monetary compensation of $20.

Stimuli and Tasks

The stimuli, design, and procedure were parallel to Yang, et al. (under review). A set of 120 two-sentence passages (mean number of words = 28) modeled after the ones shown in Table 1 was used for the experiment. The first sentence of each passage established differing degrees of antecedent referent information. The critical word was always the initial content word of the second sentence, and the first word of the second sentence was typically a determiner, such as "The".[1] Four versions of each experimental passage were constructed by varying the Integration Type (explicit, paraphrase, inference, and baseline). The four versions of the 120 experimental passages were balanced by a Latin-square that assigned each version of a passage to one of four sets of materials. Each participant received 30 passages from each of the four conditions without repetition of any passage. In this way, each participant read one condition of the experimental passage only once, and across the participants, each version of all of 120 passages was represented equally.

Design and Procedure

The EEG was recorded as the participants read each two-sentence passage for comprehension. Each word of the passage was presented on the center of the computer screen for a duration of 300 ms with a stimulus-onset asynchrony (SOA) of 600 ms. The sentence was preceded by a fixation mark to orient participants' attention to the center of the computer screen before the trial started. Participants initiated a trial by pressing the space bar. Each word occupied a 2 cm height × 4 cm width area with white text on a black background. A true-or-false comprehension task appeared after 1/4 of the trials, with feedback on correctness provided. Half of the comprehension questions were true and half were false. Correctly answering the questions required understanding of the meaning of the passage. To exemplify, the comprehension question for the sample passage of Table 1 was the following: *True or False*: *The bomb exploded when it hit the ground*. The correct answer for this example question is true for the explicit, paraphrase, and inference conditions; and false for the baseline condition.

To reduce the artifacts of eye movements and eye blinking, participants were instructed to remain as still as possible with their eyes on the center of the computer screen until the final word of the passage. They were requested to refrain from blinking as much as possible when stimuli were presented, but were told that they could rest before initiating the next trial. Participants were tested in a series of two experimental blocks of approximately 60 to 90 minutes. They were tested individually in a quiet and comfortable environment, while the experimenter monitored the ERP recordings and experiment events in an adjacent room. Participants were seated in an adjustable chair so that subjects' eyes were about 55 cm from the center of the 15-inch CRT monitor screen and that subject's gaze was level with the center of the computer screen.

Apparatus

The electroencephalogram (EEG) was recorded using a 128 channel Electrical Geodesics system (Tucker, 1993) consisting of Geodesic Sensor Net electrodes, Netamps, and Netstation software running on an Apple Macintosh 1000 MHz. PowerPC G4 class computer with Mac OS 9.2.2. (Apple Computer, Cupertino, CA). The data were recomputed off-line against the average reference, the vertex (Lehmann & Skrandies, 1980). Impedances were maintained below 50 kΩ, an acceptable level for the electrodes and amplifier used (Ferree, Luu, Russell & Tucker, 2001; Tucker, 1993). The EEG was amplified and analog filtered with
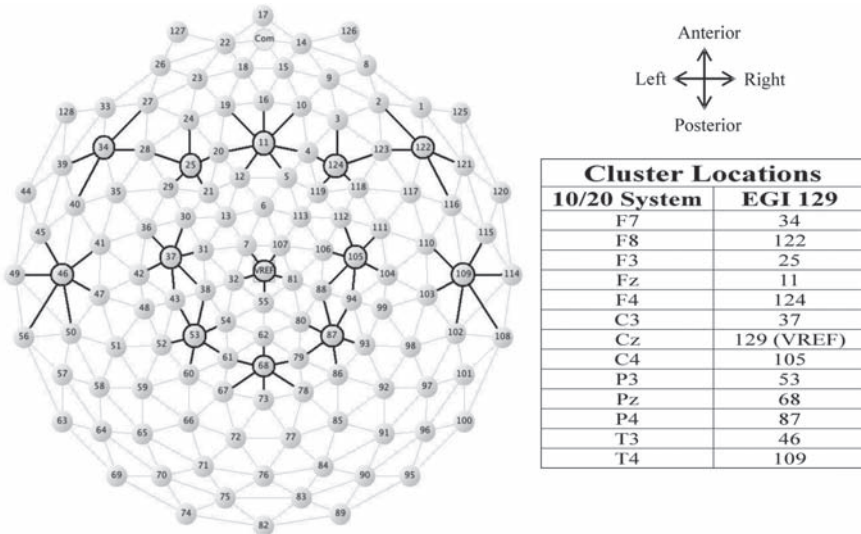
Cluster Locations	
10/20 System	**EGI 129**
F7	34
F8	122
F3	25
Fz	11
F4	124
C3	37
Cz	129 (VREF)
C4	105
P3	53
Pz	68
P4	87
T3	46
T4	109

Figure 1. The 13 clusters formed around the mostly commonly used electrode sites in the international 10–20 system (F7-F8, F3-Fz-F4, C3-Cz-C4, P3-Pz-P4, and T3-T4) for the ANOVA of the ERP analysis.

.1 Hz to 100 Hz bandpass filters, referenced to the vertex, and a 60 Hz notch filters then digitized at 250 Hz. Six eye channels were used to monitor the trials for eye movement and blinks. The EEG signals were recorded continuously at 500 Hz by the Net Station with a 12 bit A/D converter. The EGI Net Station also recorded all event onset times and accuracy for later analysis. Instruction and visual stimuli were presented on a 15-inch CRT monitor working at 60 Hz refresh rate. The experimental trials were controlled by Eprime, a commercial software (Psychology Software Incorporation, Pittsburgh, Pennsylvania), designed to present the trials and to record relevant trial information. Eprime also sent event information to the electroencephalogram (EEG) recording system (Net Station, Electrical Geodesics Inc., Eugene, Oregon).

ERP Recordings

Raw EEG data were recorded continuously, referenced to the vertex, at 1000 samples per second. The EEG data were segmented off-line into 1000 ms epochs spanning 200 ms pre-stimulus to 800 ms post-stimulus for the critical word to examine the ERPs patterns as a function of text integration. There was a total of 30 possible trials per participant per condition. Data were digitally screened for artifact (eye blinks or movements, subject movement, or transient

electronic artifact) and contaminated trials were removed. Remaining data were sorted by condition and averaged to create the ERPs. Averaged ERP data were digitally filtered at 30 Hz lowpass to remove residual high-frequency noise, baseline corrected over the 200 ms pre-stimulus period, and re-referenced into an average reference frame to remove topographic bias due to choice of reference site (Dien, 1998). The subject-averaged ERPs were averaged together to produce the mean waveform across subjects, the grand average waveform for each condition, which was used in subsequent topographic plots and statistical analyses of brain waveforms. Figure 1 shows the 128 electrodes of the EGI Geodesic Sensor Net and 13 clusters of electrodes selected for analysis based on 13 electrodes of the 10/20 system, which is commonly used in ERP research. Grand average waveforms were computed for these 13 clusters as shown in Figure 2. Figure 3 shows topographic maps of the voltage field across the surface of the scalp, created by interpolating from the recorded values using spherical splines (Perrin, Pernier Bertrand & Echallier, 1989).

ERP Data Analysis Strategies

Our strategy was to acquire a converging view of the ERP results through the use of multiple complementary analyses. These include a temporal Principal Component Analysis (PCA), tests of mean amplitude differences (brainwave analysis), and topographic voltage maps of difference waves.

Topographic maps. The topographic maps (Figure 3) are especially useful for showing the spatiotemporal dynamics of word processing over time. These topographic maps, together with a temporal PCA, can help to identify the ERP components for further brainwave analysis based on mean amplitudes recorded from a subset of the electrodes.

PCA. A temporal PCA was conducted on subject averages, based on 375 2-ms samples (0–750 ms post-stimulus onset). PCA extracts independent components to allow theory-free, data-driven ERP patterns to emerge based on the covariance matrix of all sample points. Each temporal factor can be considered to represent a particular pattern of temporal activity (e.g. a time window) that may be associated with a mental operation that occurs during word processing. In this way, PCA provides a quantitative identification of a specific time window of related cognitive activity. The PCA data consisted of the ERP averages at each electrode site in all the experimental conditions for each subject. The data matrix input of this temporal PCA consisted of 375 time points by 9,288 observations (129 electrode sites * 18 subjects * 4 experimental conditions). The PCA used the covariance matrix with Promax rotation (Chapman & McCrary,

1995; Dien & Frishkoff, 2005; van Boxtel, 1998). ANOVAs were conducted on the component score of each component derived from PCA.

Mean Amplitude ANOVAs. The brainwave analysis used two repeated-measures ANOVAs. One tested the ERPs for the three medial electrode sites (Fz, Cz, and Pz) and the other tested ten lateral electrode sites (F7–F8, F3–F4, C3–C4, P3–P4, and T3–T4). These thirteen electrode regions were selected because they provide sufficient coverage to examine the related ERP shifts to the effects of text integration during language comprehension.

For the brainwave analysis, ANOVAs were conducted on the mean voltage amplitudes extracted from the related timing windows of each prominent ERP component identified from the PCA. For both PCA and brainwave ANOVAs there were two within-subject factors: Integration Type (explicit, paraphrase, inference, and baseline) and Electrodes (3 for medial sites ANOVA and 5 pairs for lateral sites ANOVA) and the ANOVA of lateral sites includes an additional within-subject factor of Hemisphere (Left vs. Right). All probability values reported for effects with more than two degrees of freedom were adjusted with the conservative Geisser-Greenhouse correction for deviations from sphericity in the data.

Results

Behavioural Results

The mean accuracy of responses to the comprehension questions that followed some of the passages was as follows: explicit (84%), paraphrase (92%), inference (90%), and baseline (68%). Table 2 shows these results along with results for skilled comprehenders from Yang, et al. (under review) for comparison. Because the critical word of the baseline condition was constructed to be unrelated to a sensible integration during the immediate integration process, the baseline condition is excluded in further behavioural analysis. A repeated measure ANOVA indicated a significant effect of Integration Type ($F(2, 34) = 4.55$, $p < 0.05$, $MSE = 62.4$). According to pairwise comparisons, the explicit condition had lower accuracy than the paraphrase ($t(17) = 2.55$, $p < 0.05$), and the inference conditions ($t(17) = 2.17$, $p < 0.05$), which did not differ from each other. By comparison, the pattern for the skilled readers of Yang, et al. (under review) showed higher comprehension performance for the paraphrase condition than either the inference or the explicit condition. Thus, skilled comprehenders performed better in the paraphrase condition than in other conditions, whereas

Table 2.   Percent correct responses to comprehension questions. Conditions with "*" are those which are significantly different from other conditions. Data for skilled readers are from Yang, et al. (under review).

	Explicit	Paraphrase	Inference	AVERAGE
Skilled	91	96 *	89	92
Less skilled	84 *	92	90	88

the less skilled comprehenders performed worse in the explicit condition than the other conditions. A caution in interpreting these data is that because the comprehension questions were asked for only 1 of 4 passages and based on two-choice (true/false) responding, the percent correct may not be a particularly robust indicator of comprehension.

ERP Results

Figure 2 plots the grand average forms for each condition for each of the thirteen clusters identified in Figure 1. For comparison, Figure 2 also shows the grand average for the vertex cluster (Cz) for the skilled comprehenders of Yang et al. (under review), which reveals a clear reduction in the N400 for integration conditions (and a 300 ms positivity for the paraphrase condition) relative to baseline that is not seen in the corresponding electrode cluster for the less skilled readers.

To better characterize the spatiotemporal dynamics of the integration processing of the critical word, Figure 3 (upper part) shows the top view of topographic maps of the voltage field distributions of difference waves, for which the baseline condition is subtracted from each experimental condition. The baseline condition produced the expected posterior negativity (N400), along with a reciprocal anterior positivity, at the critical word.

The experimental conditions show growth of posterior positivities and anterior negativities, reflecting their relative ease of integration. Differences in two time frames are visible in Figure 3. A middle time frame from ~300 ms to ~400 ms shows a reciprocal negativity-positivity pattern salient along the anterior-posterior axis for the explicit and inference conditions. No apparent processing difference for the paraphrase condition was observed at this anterior-posterior axis. A later time frame from ~400 ms to 750 ms indicates a more prominent pattern of reciprocal negativity-positivity along the anterior-posterior axis for both the explicit and inference conditions. Importantly, the paraphrase condition showed a different pattern at this late time window. One

can see (Figure 3) a developing posterior positivity and a developing right-anterior negativity for the processing of paraphrased words.
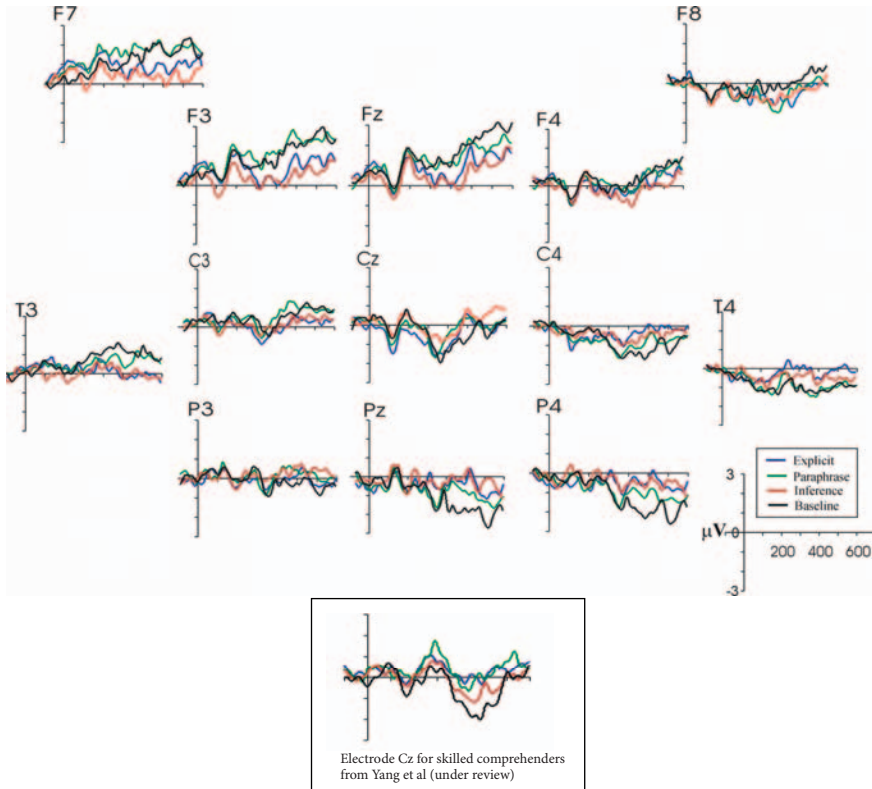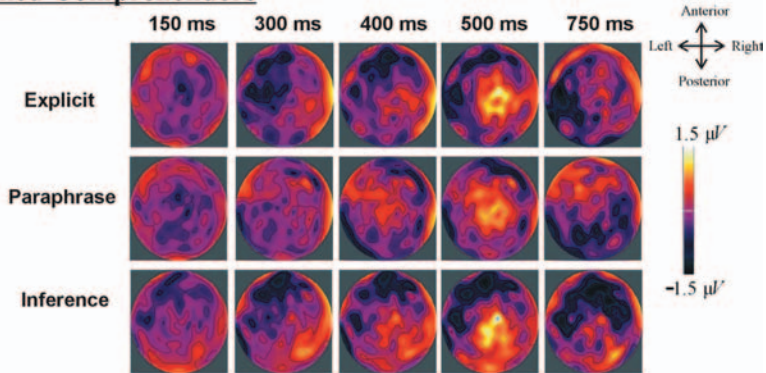

Electrode Cz for skilled comprehenders
from Yang et al (under review)

Figure 2. Grand average reference ERPs recorded from the onset of the critical word for 13 clusters plotted for each of the four conditions of the experiment. The enclosed graph at the bottom shows the grand average ERP for the skilled readers of Yang et al. (under review) at cluster Cz. The Cz cluster shows an N400 reduction effect for the explicit and paraphrase conditions for the skilled comprehenders but not for the less skilled comprehenders. For the less skilled group, the ERP differences between experimental conditions were most salient in two scalp regions: the anterior sites of F3 and Fz, and the right-dominance posterior sites of C4 and P4 that included Pz. The N400 reduction as a function of text integration occurred at the right-dominance posterior sites. The temporal sites showed ERPs related to these two major regions, but the time course differed.

PCA Components

The PCA extracted 6 major temporal factors (eigen value> 1.5) that accounted for 88% of the total variance in the data from the 375 variables corresponding to each 2 ms ERP time frame (See Figure 4). Because the fourth through the sixth components (accounting for 7% of the total variance) have early peaks and were not modulated by the Integration Type factor, they are not of major interest to our comprehension questions. They appear to reflect exogenous

Figure 3. Comparison of skilled and less skilled comprehenders in ERP spatial-temporal dynamics. Time-sequenced topographic maps show the voltage field distributions of difference waves, with the baseline condition subtracted from each experimental condition. The horizontal axis shows the temporal sequences of these distributions from 150 ms to 750 ms. Positivite voltage shifts relative to the baseline condition show as lighter areas, with negativities darker. The development of the positivities appears to reflect the unfolding of word identification and word-to-text integration. The slower developing positivites for less skilled comprehenders appear to reflect a slower and less robust integration process.

Less Skilled Comprehenders
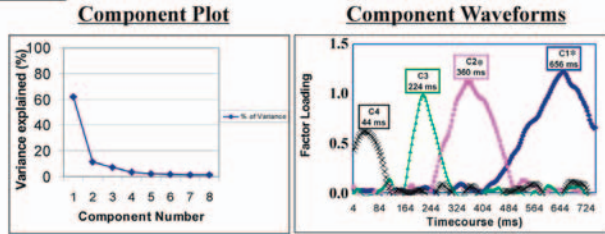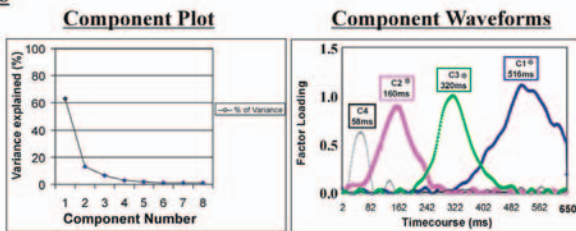

Skilled Comprehenders

Figure 4. Scree plot and component waveforms of extracted PCA components for the less skilled comprehenders of the current study and the skilled comprehenders of Yang et al. (under review). For each group, the left panel shows the scree plot of the first 8 components extracted from following promax rotation. For the component waveforms, the first four components (component 1, C1; component 2, C2; etc.) and the time points of their peak loadings are shown. Components with asterisks were sensitive to experimental conditions. Note that "component" here refers to a PCA component, not to voltage shifts.

components (C 5, 132 ms; and C 6, 164 ms) related to visual processing and equipment artifact (C 4). Accordingly, the first three PCA components (80% of total explained variance) were selected for further statistical analysis (Note that "component" in this section refers to PCA component, not to voltage shifts).

Component 1 (late-latency component, 62% explained variance) is a slow wave component, rising from approximately 500 ms and reaching a peak at 656 ms. This peak matched the enhanced negative shift in the EEG data at C4, Pz, and P4 (Figure 2, & Figure 3). The average component scores were negative for parietal sites and positive for frontal, central and temporal sites. Statistical tests indicated that Component 1 was modulated by Integration Type at both medial and lateral sites. There was significant interactions of Integration Type x Electrodes for the medial site ANOVA ($F(6, 102) = 4.37$, $p < 0.01$) as well as Integration Type x Electrodes ($F(12, 204) = 3.58$, $p < 0.01$), and Integration Type x Hemisphere ($F(3, 51) = 3.42$, $p < 0.05$) for the lateral sites ANOVA. The baseline condition elicited greater negativity than the explicit and inference conditions

at the parietal site (Pz,& P4), but greater positivity than the explicit and inference conditions at the medial frontal sites (F3, & Fz). In addition, the positivity elicited by the paraphrase condition was greater than the explicit and inference conditions only at the left anterior site (F3).

Component 2, (mid-latency component, 11% explained variance) rose from ~300 ms with a peak at 360 ms. The average component scores were positive at the frontal and temporal sites, and negative at the central, and parietal sties. The peak and time course matched the negative shift in the EEG data at the central and parietal sites. ANOVAs indicated a significant interaction effect of Integration Type x Electrodes (F(6, 102) = 2.39, $p < 0.05$) only at the medial site. The baseline condition elicited greater negativity than the explicit condition at the medial parietal site (Pz), but greater positivity than the explicit and inference conditions at the medial frontal site (Fz).

Component 3, (early-latency component, 7% explained variance) rose from ~180 ms, with a peak at 220 ms. The ERP waveform of Figure 2 indicates that most electrodes sites show negative polarity at this peak except for the frontal sites (F3, Fz, and F4). The average component scores were positive for frontal site, but negative at the central, parietal and temporal sites. ANOVAs showed that this component is not modulated by the Integration Type. Based on the latency, polarity and scalp distribution, this component indicated early lexico-semantic level of word processing.

Thus, the PCA identified temporal components that are consistent with the topographic analyses. It further identified components that relate to semantic-conceptual and integration processing (components 1, & 2), separable from those components that relate to earlier stages of word identification processing (component 3).

Although we postpone detailed comparison of skilled and less skilled comprehenders until the discussion, we point out here that Figure 4 shows both a parallelism and some divergence between the components for skilled and less skilled comprehenders. Components 1–3, all reflecting cognitive components, are aligned over similar temporal points for the two groups. However, the components peaks are noticeably later for the less skilled comprehenders, consistent with generally slower word identification and integration processes.

Mean Amplitude Analysis

To further characterize the temporal and spatial patterns in the data, we carried out analyses of the brainwave forms that correspond to the two PCA components (mid- and late latency) that were affected by the integration conditions.

To effectively increase the accuracy of the average-reference derivation (Dien, 1998; Tucker, Liotti, Potts, Russell & Posner, 1994), we divided the electrode sites into 13 regions (3 lateral (left/midline/right) and 5 lobes (prefrontal/frontal/central/posterior/temporal)) based on the 10/20 system (Figure1: F7–F8, F3–Fz–F4, C3–Cz–C4, P3–Pz–P4, and T3–T4). In each region, we averaged the mean voltage amplitude of surrounding channels. This procedure provides more reliable sampling of the activity within any region (Dien & Santuzzi, 2005) thus reducing spurious interactions obtained with single electrode locations.

Middle (280 to 460 ms) and late (510 to 750 ms) time windows were tested in two separate repeated-measures ANOVAs. The time windows were defined by converging sources: the components of PCA analysis (Figure 4), the grand average reference ERP (Figure 2), and the spatiotemporal dynamics of topographic voltage distribution of difference waves (Figure 3). The selected time windows not only were consistent with the time course of a PCA component waveform but also maintained the greatest amplitude divergence among experimental conditions. The two ANOVAs, one for lateral sites and one for medial sites, were conducted on mean voltage amplitudes extracted from each time window.

Middle Time Frame (from 280 to 460 ms)

Only the midline analysis showed effects of the experimental variables in this time frame. Both the explicit and inference conditions induced prominent brainwave differences, relative to the baseline condition. The midline ANOVA confirmed these differences in a significant interaction of Integration Type x Electrodes (F(6, 102) = 3.08, $p < 0.05$). Post-hoc contrasts showed significant differences between the explicit and baseline conditions ($t(17) = -2.51, p < 0.05$), and between the inference and baseline conditions ($t(17) = -2.88, p < 0.01$) at the medial frontal site (Fz). At the medial parietal site (Pz), these significant patterns persist, but in an opposite direction (the explicit vs. baseline condition ($t(17) = 2.15, p < 0.05$), and the inference vs. baseline conditions ($t(17) = 2.21, p < 0.05$)). No significant differences of experimental conditions were found at the medial central site (Cz).

Late Time Frame (from 510 to 750 ms)

For the late 510 to 750 ms time window, both the midline and lateral analyses showed a significant interaction of Integration Type x Electrodes (F(6, 102) = 4.76, $p < 0.01$; F(12, 204) = 3.18, $p < 0.05$). A significant interaction

of Integration Type x Hemisphere (F(3, 51) = 3.22, *p < 0.05*) was also found in the lateral analysis. For the midline region, post-hoc contrasts indicated that the baseline condition was more positive than the other conditions (vs. explicit, $p < 0.05$; vs. paraphrase, $p < 0.05$; vs. inference, $p < 0.005$) and that the paraphrase condition was more positive than the inference condition ($p < 0.05$) at the anterior site (Fz). These patterns were opposite at the posterior site (Pz), with the baseline condition more negative than the other conditions. The difference between the baseline condition and the inference condition was reliable ($p < 0.05$), whereas differences between the baseline and the other conditions were of marginal reliability (vs. explicit, $p = 0.06$; vs. paraphrase, $p = 0.08$). No other processing difference was found at the medial central (Cz).

Similar patterns were observed in the interaction of Integration Type x Electrodes for the lateral analysis. The anterior region (prefrontal and frontal) indicated significant differences between paraphrase vs. inference conditions ($p < 0.05$) and between inference vs. baseline conditions ($p < 0.01$). Again, the parietal site showed significant differences between inference and baseline conditions, but in an opposite direction ($p < 0.005$). The differences between paraphrase and baseline conditions reached only marginal significance ($p = 0.07$).

The Integration Type x Hemisphere for the lateral analysis reflected a left hemisphere difference between both paraphrase and baseline conditions vs. explicit and inference conditions. Post hoc contrasts of these differences were significant, except that the explicit-baseline comparison was marginal ($p = .08$). No other contrast was found significant. The right hemisphere did not show experimental effects.

Discussion

The ERP results provide a convergent picture of the word-by-word integration processes that occur across sentence boundaries. Consistent with our previous study of skilled comprehenders (Yang, et al., under review), less skilled comprehenders showed evidence for integration across a sentence boundary when they read the first words of a sentence. However, the less skilled comprehenders of this study show a different pattern from those observed for skilled comprehenders by Yang, et al. In the following discussion, we summarize these differences and examine their implications.

ERP Comparisons with Skilled Comprehenders

To characterize these skill differences, we briefly review the logic of our experimental design. Our test of word-to-text integration depends on finding ERP indicators that a reader has related the first content word of a sentence to a referent that may have been established by the first sentence. The accessibility of the potential referent depended on the first sentence, as varied across our four experimental conditions. The word-to-text integration could be lexically-based (explicit condition), semantically and conceptually-based (paraphrase condition), or situationally based (predictive inference). The integration of the baseline condition was the most difficult, because the critical word of the second sentence had no apparent antecedent in the wording of the first sentence, nor in the situation suggested by the first sentence. The ERP effects on the critical word in the second sentence depended on these integration conditions. In particular, the N400 (along with converging evidence from PCA and topographic analyses) marks the difficulty of the integration. Indeed, the N400 was largest for the baseline condition, which was expected to be the least favourable for integration.

To compare the integration results for skilled and less skilled comprehenders, we summarize the ERP results for the skilled comprehenders of Yang, et al. (under review) in the topographic maps of Figure 3 (lower part). Skilled comprehenders show the changes in amplitude for each experimental condition with the baseline amplitudes subtracted. Negativity shows as black and lightens with positivity. One can see ERP positivities emerge by ~150 ms and continue through two later shifts identified as a P300 and N400. The early (from ~150 ms to ~250 ms) developing posterior positivity (relative to baseline condition), which was observed for explicit and paraphrase conditions, suggests lexico-semantic processing that may be the first phase of integration. The following P300, observed only for the paraphrase condition at the midline site, suggests a semantic-conceptual process triggered by a paraphrase. The major N400 reduction for the explicit and paraphrase conditions suggests that this standard indicator of word integration was affected by the word's meaning in relation to the preceding context. By contrast, the enhanced N400 for the inference and baseline conditions suggests more integration effort was required for situationally-based rather than word-based integration.

The results for the less skilled comprehenders of the present study show two major departures from this pattern. First, the less skilled comprehenders showed no effects (prior to 300 ms) that would reflect an early phase of word processing that is sensitive to integration opportunities. Second, they showed

a two-phase integration process after ~300 ms, with a middle peak around 360 ms and a late peak around 656 ms, which were identified by PCA. Both the PCA and mean amplitude analysis indicate qualitative differences between these two shifts. Figure 4 indicates the PCA patterns for both less skilled (upper part) and skilled comprehenders.[2]

The middle component indicates that both the explicit and inference conditions differed from the baseline condition at the anterior-posterior axis sites (Fz-Pz). At the posterior site, the reduced N400 of the explicit condition suggests integration by word repetition. Additional differences between experimental conditions were observed at the late time frame, showing a later phase of word-to-text integration that follows word identification. The anterior sites show multiple differences between conditions, with the baseline condition reliably different from all other conditions and with the paraphrase condition different from the inference condition. In contrast, the posterior site observed reliable differences only between the inference and the baseline conditions. This contrasts with skilled comprehenders, who showed high differentiation of the N400 only in the central-parietal areas.

An interesting characteristic of the late-time component is that the paraphrase condition produced differences from the inference and baseline conditions in anterior regions. This separation of word-meaning-based from situationally-based integration suggests that, like skilled comprehenders, less skilled comprehenders take advantage of word meanings in integrating a word with a text. However, this paraphrase effect emerges later for the less skilled readers, suggesting a slower activation of the word's meaning or a slower selection of the meaning features that would provide integration with the given context. In a word, the lexical processing that should produce context-relevant meanings is sluggish for these less skilled comprehenders.

Underlying this general characterization is a trend for the paraphrase condition to begin its separation from other conditions as early as 200 ms, although it is not until about 510 ms that this separation becomes statistically reliable(see Figure 2). The non-reliable early trend may reflect differences among readers and words that affect variance estimates. More generally, the anterior-dominant topographic pattern for the paraphrase effect shown by less skilled comprehenders points to the possibility that they rely on text memory processes to support integration, compensating for their inefficient lexical processing skill.

Perfetti and Hart (2002) argued that the quality of a reader's lexical representation affects both the speed of word retrieval and the meaning selection of ambiguous words. A consequence of a lower quality representation — one

whose orthographic, phonological, and semantic components are incompletely specified and/or less well integrated — is that reading the word activates related concepts (e.g. paraphrases) more slowly. In reading for comprehension, as in the present study, readers integrate the critical word rapidly with the preceding context by identifying the word's form and its meaning and selecting a meaning relevant to the contents of an active text memory. These word processes are less effective or less efficient for the less skilled comprehender. The ERP shifts track these processes, indicating the less effective use of word repetition and especially meaning-related paraphrase words, and further suggesting that cognitive processes associated with anterior regions are needed to compensate for less effective lexical processes. If this characterization is correct, the quality of the reader's lexical representation affects both the initial process of word identification and the cognitive dynamics in the processing of word-to-text integration.

The comprehension-question results allow a tentative examination of the ERP results in relation to the comprehension outcomes. Here, the behaviour data, consistent with Yang, et al. (under review), indicates a dissociation between immediate integration and final comprehension. In the present study, the major ERP analyses suggested that integration was facilitated by both text-level (explicit) and situation level (inference) effects for less skilled comprehenders. In contrast, the comprehension accuracy suggested that both the paraphrase and inference conditions facilitated comprehension whereas the explicit condition did not. Thus, for both skilled and less skilled readers, the explicit reinstatement of a word promoted immediate integration of the word with the preceding text, but it did not actually promote comprehension better than the paraphrase and inference conditions. As can be seen in Table 2, skilled comprehenders had 4% overall higher accuracy to the comprehension questions than did the less skilled comprehenders, a reliable difference according to an across study analysis of variance.

The advantage of paraphrase over a word repetition may be traced to a more active vs. a more passive integration process (Schmalhofer & Perfetti, in press; Yang et al., under review). For the explicit word integration, a process of passive memory resonance is sufficient to integrate the repeated word with its antecedent, a word-to-word process. However, this level of argument overlap (Kintsch, 1988) based on surface features may occasionally fail to create word-to-referent text integration. On the other hand, the semantic-conceptual processing induced by a paraphrase leads to an active process that links the new word — the paraphrase — to a referent established in the situation model.

In this picture, comprehension differences can arise from lexical processing skills. The context-sensitive lexico-semantic and conceptual processes of skilled comprehenders emerge as early at 150 ms (Yang et al., under review). These early processes allow the later integration processes at 300–400 ms (observed in posterior electrodes) that link the word to the reader's text representation based on the word's meaning. In contrast, with inefficient lexical processing skill, less skilled comprehenders do not show these early lexico-semantic processes until around 300 ms. And at that point, the ERP record shows an anterior-dominant distribution (not seen in skilled comprehenders) that may indicate additional memory processes rather than simple word integration.

Conclusions and Implications for Comprehension Skill

Our results have several implications for the basis of individual differences in comprehension. One is that comprehension weaknesses assessed by standardized tests arise in word-by-word sentence reading. Our texts were short and simple and did not afford opportunity for higher level text structure factors to emerge. The differences in ERP patterns between this sample of less skilled comprehenders and the more skilled comprehenders of Yang, et al. (under review) were notable and were especially dramatic in the temporal dynamics revealed by the topographic maps.

One rather important difference concerns the inference condition, in which easy integration is possible only if the reader had made a forward inference during the first sentence. Although less skilled readers are often found to have problems in making inferences (e.g., Cain & Oakhill, 1999), our findings give no support to this possibility in our adult sample. Indeed, we found that less skilled comprehenders showed evidence for integration on the critical word when its referent could have been established by a forward inference in the preceding sentence. By contrast, although the trend was in the same direction, the skilled comprehenders of Yang, et al. (under review) did not show this inference-based integration effect reliably. Thus, less skilled comprehenders actually seemed to make forward inferences more reliably than did skilled comprehenders. We note that the ERP study by St. George, et al. (1997) reached the opposite conclusion based on N400 reductions measured on short passages. There were material and group definition differences between the St. George, et al.'s study and ours that could have made a difference. Perhaps a more direct factor is that our measure was identical across conditions, whereas the St. George, et al.'s study reported averages taken across the whole sentence. In any

case, the word-by-word study of integration has the potential to shed light on this issue in a way that studies using less fine grain measures cannot.

The early use of paraphrase by skilled comprehenders and the very slow and weak use of this information by less skilled comprehenders is another conclusion to emphasize. Whereas skilled comprehenders showed the paraphrase effect as a P300 as well as the more typical N400, less skilled comprehenders did not show an N400 effect for paraphrase. A later effect did emerge, but its lateness and its distribution suggests its interpretation is not the same as the N400 reduction. This slowness of interpreting a new word as related to a previous word in the text could be an important general source of comprehension problems.

Despite the skill group differences in patterns, we do not take these differences to be qualitative. The emergence of ERP positivities associated with ease of integration was slower for less skilled comprehenders rather than completely absent. The picture of word-to-text integration for the less skilled comprehender is, as our title suggests, "sluggish".

Acknowledgements

Notes

1. The full set of stimulus materials (120 passages) can be viewed on the web-site: http://my.pitt.edu/~cyang/integration_paper.pdf. The first content words of the second sentences have variability in form class, but grammatically they all constituted the initial noun phrase of the second sentence.

2. The PCA figure rescales the Y axis of factor loading from Yang et al. (under review) to make it comparable with the PCA of the current study.

References

Baker, L. (1982). An evaluation of the role of metacognitive deficits in learning disabilities. *Topic in Learning and Learning Disabilities, 2*, 27–35.

Cain, K. (1996). Story knowledge and comprehension skill. In C. Cornoldi. and J. Oakhill (Eds.), *Reading comprehension difficulties* (pp. 167–192). Mahwah, NJ: Erlbaum.

Cain, K. and J. V. Oakhill (1999). Inference ability and its relation to comprehension failure in young children. *Reading and Writing, 11*, 489–503.

Chapman, R. M. and J. W. McCrary (1995). EP component identification and measurement by principal components analysis. *Brain and Cognition, 27,* 288–310.

Cornoldi, C. and J. Oakhill (1996). In C. Cornoldi and J. Oakhill (Eds.), *Reading comprehension difficulties*. Mahwah, NJ: Erlbaum.

Dien, J. (1998). Issues in the application of the average reference: review, critique, and recommendations. *Behavioral Research Methods, Instruments and Computers, 30,* 34–43.

Dien, J. and G. A. Frishkoff (in press). Introduction to Principal components analysis of event related potentials. In T. Handy (Ed.), *Event-Related Potentials: A methods handbook.* Cambridge, MA: MIT Press.

Dien, J. and A. M. Santuzzi (in press). Application of repeated measures ANOVA to high density ERP datasets: A review and tutorial. In T. Handy (Ed), *Event-Related Potentials: A methods handbook.* Cambridge, Mass: MIT Press.

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

Ferree, T. C., Luu, P., Russell, G. S. and D. M. Tucker (2001). Scalp Electrode Impedance, infection risk, and EEG data quality. *Journal of Clinical Neurophysiology, 112(3),* 536–544.

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.

Karis, D., Fabiani, M. and E. Donchin (1984). P300 and memory: Individual differences and the Von Restorff effect. *Cognitive Psychology, 16*, 177–216.

King, J. W. and M. Kutas (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading, *Journal of Cognitive Neuroscience, 7(3),* 376–395.

Kintsch, W. (1988). The role of knowledge in discourse comprehension. A construction-integration model. *Psychological Review, 95*, 163–182.

Kintsch, W. and van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85,* 363–394.

Kluender, R. and M. Kutas (1993). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies, *Journal of Cognitive Neuroscience, 5(2),* 196–214.

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

Kutas, M. and J. W King (1996). The potentials for basic sentence processing: Differentiating integrative processes. In I. Ikeda and J. L. McClelland (Eds.), *Attention and performance* (pp. 501–46). Cambridge, MA: MIT Press.

Landi, N. (2005). *Behavioral and electrophysiological investigations of semantic processing in skilled and less-skilled comprehenders*. Doctoral dissertation, University of Pittsburgh. http://etd.library.pitt.edu/ETD/available/etd–08182005–133452/

Lehmann, D. and W. Skrandies (1980). Reference-free identification of components of checherboard-evoked multi-channels potential fields. *Electroencephalography and Clinical Neurophysiology, 48,* 609–621.

Long, D. and J. Golding (1993). Superordinate goal inferences: Are they automatically generated during comprehension? *Discourse Processes, 16,* 55–74.

Long, D. L., Oppy, B. J. and M. R. Seely (1997). Individual differences in reader's sentence- and text-level representations. *Journal of Memory and Language, 36,* 129–145.

Nation, K. (2005). Children's reading comprehension difficulties. In M. J. Snowling and C. Hulme (Eds.), *The science of reading: A handbook* (pp. 248–265). Oxford: Blackwell.

Nation, K. and M. J. Snowling (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. and A. Garnham (1988). *Becoming a skilled reader*. New York: Basil Blackwell.

Perfetti, C. A., Landi, N. and J. Oakhill (2005). The acquisition of reading comprehension skill. In M. J. Snowling and C. Hulme (Eds.). *The science of reading: A handbook* (pp. 227–247). Oxford: Blackwell.

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

Perfetti, C. A. and L. Hart (2001). The lexical basis of comprehension skill. In D. Corfien (Ed.), *On the consequences of meaning selection* (pp. 67–86). Washington, DC: American Psychological Association.

Perfetti, C. A. (1999). Comprehending written language: A blueprint of the reader. In C. Brown and P. Hagoort (Eds.), *The neurocognition of language* (pp. 167–208). Oxford University Press.

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

Perrin, F., Pernier, J., Bertrand, O. and J. F. Echallier (1989). Spherical splines for scalp potentials and current density mapping. *Electroencephalography and Clinical Neurophysiology, 72*, 184–187.

Rosler, F., Heil, M. and E. Hennighausen (1995). Exploring memory functions by means of brain electrical topography: A review. *Brain Topography, 7*, 301–313.

Ruchkin, D. S., Johnson, R., Mahaffey, D. and S. Sutton (1988). Toward a functional categorization of slow waves. *Psychophysiology, 25,* 339–353.

Schmalhofer, F. and C. A. Perfetti (Eds.). (in press). *Higher level language processes in the brain: Inference and comprehension processes*. Mahwah, NJ: Erlbaum.

Singer, M. and M. Halldorson (1996). Constructing and validating motive bridging inferences. *Cognitive Psychology, 30,* 1–38.

St. George, M., Mannes, S. and J. E. Hoffman (1997). Individual differences in inference generation: An ERP analysis. *Journal of Cognitive Neuroscience, 9*(6), 776–787.

Tucker, D. M. (1993). Spatial sampling of head electrical fields: The geodesic sensor net. *Electroencephalography and Clinical Neurophysiology, 87*, 154–163.

Tucker, D. M., Liotti, M., Potts, G. F., Russell, G. S., and Posner, M.I. (1994). Spatiotemporal analysis of brain electrical fields. *Human Brain Mapping, 1*, 134–152.

van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V, Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(3), 443–467.

van Berkum, J. J. A., Zwitserlood, P., Brown, C. M. and P. Hagoort (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 Berkum, J. J. A., Hagoort, P., and C. M. Brown (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience, 11*(6), 657–671.

van Boxtel, G. J. M. (1998). Computational and statistical methods for analyzing event-related potential data. *Behavioral Research Methods, Instruments, and Computers, 30*, 87–102.

Van Petten, C., Weckerly, J., McIsaac, H. K. and M. Kutas (1997). Working memory capacity dissociates lexical and sentential context effects. *Psychological Science, 8*, 238–42.

Van Petten, C. and M. Kutas (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory and Cognition 18*, 380–393.

Vos, S. H. and A. D. Friederici (2003). Intersentential syntactic context effects on comprehension: the role of working memory. *Cognitive Brain Research, 16,* 111–122.

Yang, C. L., Perfetti, C. A. and F. Schmalhofer (under review). ERP indicators of text integration across sentence boundaries.

Author's addresses:

Chin Lung Yang
Learning Research and Development Center
University of Pittsburgh
3939 O'Hara Street, Room# 413
Pittsburgh, PA 15260
cyang@pitt.edu