Semantic Cortical Activation in Dyslexic Readers

Päivi Helenius and Riitta Salmelin
Helsinki University of Technology

Elisabet Service
University of Helsinki

John F. Connolly
Dalhousie University

Abstract

The combined temporal and spatial resolution of MEG (magnetoencephalography) was used to study whether the same brain areas are similarly engaged in reading comprehension in normal and developmentally dyslexic adults. To extract a semantically sensitive stage of brain activation we manipulated the appropriateness of sentence-ending words to the preceding sentence context. Sentences, presented visually one word at a time, either ended with a word that was (1) expected, (2) semantically appropriate but unexpected, (3) semantically anomalous but sharing the initial letters with the expected word, or (4) both semantically and orthographically inappropriate to the sentence context. In both subject groups all but the highly expected sentence endings evoked strong cortical responses, localized most consistently in the left superior temporal cortex, although additional sources were occasionally found in more posterior parietal and temporal areas and in the right hemisphere. Thus, no significant differences were found in the spatial distribution of brain areas involved in semantic processing between fluent and dyslexic readers. However, both timing and strength of activation clearly differed between the two groups. First, activation sensitivity to word meaning within a sentence context began about 100 msec later in dyslexic than in control subjects. This is likely to result from affected presemantic processing stages in dyslexic readers. Second, the neural responses were significantly weaker in dyslexic than in control subjects, indicating involvement of a smaller or less-synchronous neural population in reading comprehension. Third, in contrast to control subjects, the dyslexic readers showed significantly weaker activation to semantically inappropriate words that began with the same letters as the most expected word than to both orthographically and semantically inappropriate sentence-ending words. Thus, word recognition by the dyslexic group seemed to be qualitatively different: Whereas control subjects perceived words as wholes, dyslexic subjects may have relied on sublexical word recognition and occasionally mistook a correctly beginning word for the one they had expected.

INTRODUCTION

The functional anatomy of impaired reading in developmental dyslexia is still poorly known. Previous imaging studies have found differences between normal and dyslexic readers in the activation of various cortical areas. However, due to limited understanding of the cortical basis of normal reading it has not been possible to fully infer the meaning of these differences. We used magnetoencephalography (MEG) to investigate simultaneously in spatial and temporal domains the activation of the semantic cortical network involved in reading sentences. Cortical activation evoked by visually presented sentence-ending words, graded with respect to their appropriateness to the preceding sentence context, was compared between reading-disabled and normally reading subjects.

Dyslexia Affects Reading Even in Adults

Reading and writing are extremely complicated skills requiring mastery and smooth cooperation of several cognitive subskills. Accordingly, in some of us, the acquisition of literacy can be seriously hampered. About 4% of children (Hulme, 1987) have difficulties in learning to read or to write, due to a condition known as developmental dyslexia. Assumed at first to be a uniquely visual problem (Morgan, 1896), dyslexia has for the past 15 years been postulated to derive from a core deficit of
phonological processing (Bradley & Bryant, 1983; Shaywitz, 1996; Wagner & Torgesen, 1987).

Phonological processing refers to the use of sound information of spoken language. While learning to read, a child must be able to detect the phonemes that constitute spoken words to understand that these separate sound units can be represented by letters or letter groups, graphemes, in written language. If phoneme segmentation is not mastered, fluent reading and writing are hard to achieve. Despite the dominating role of phonological explanations of dyslexia, problems with phonology may not be the sole cause of reading retardation. Developmental dyslexia is ultimately a heterogeneous condition (Ellis, 1985) because a deficit in one or in a combination of cognitive operations other than phonology (e.g., visual processing and lexical or semantic access) can also lead to difficulties in reading (Seymour, 1987). Following the initial problems in acquisition of literacy, dyslexic individuals remain slow and error prone in reading and writing even as adults (Scarborough, 1984).

**Imaging Studies on Reading in Normal and Dyslexic Adults**

The psychological models describing the essential operations in reading and their relative arrangement are still the subject of considerable debate (McClelland, 1987; Tousman & Inhoff, 1992; Coltheart, Curtis, Atkins, & Haller, 1993; Lukatela & Turvey, 1994a, 1994b). The questions have not been solved with imaging studies either, because the data are fairly inconclusive about the roles of the various cortical areas activated in reading tasks. The involvement of extrastriate occipital (Petersen, Fox, Posner, Mintun, & Raichle, 1988; Pugh et al., 1996) and inferior occipitotemporal areas (Halgren, Baudena, Heit, Clarke, & Marinkovic, 1994; Nobre & McCarthy, 1995; Salmelin, Service, Kiesilä, Uutela, & Salonen, 1996) in visual word form analysis has been suggested. Phonological processing in reading (i.e., grapheme-to-phoneme conversion) may be mediated by Broca's area (Pugh et al., 1996; Shaywitz, 1996) and by parieto-temporal areas (Petersen et al., 1988). Further, according to Howard et al. (1992) lexicons for written and phonological word forms might be localized to the left superior and middle temporal gyri. Semantic processing has been thought to engage the left superior/middle temporal cortex (Bavelier et al., 1997; Halgren et al., 1994; Helenius, Salmelin, Service, & Connolly, 1998; Just, Carpenter, Keller, Eddy, & Tulborn, 1996; Price, Moore, Humphreys, & Wise, 1997; Pugh et al., 1996; Vandenberghe, Price, Wise, Josephs, & Frackowiak, 1996) and its right-hemisphere homologue (Just et al., 1996). Semantic tasks also frequently activate the left inferior frontal gyrus, anterior to Broca's area (Petersen et al., 1988; Vandenberghe et al., 1996). However, this area has been suggested to control retrieval of semantic information rather than to actually store con-ceptual knowledge (Demb et al., 1995; Gabrieli, Russell, & Desmond, 1998). In addition to these lateral sites, medial temporal structures have been shown to be involved in semantic processing (Halgren et al., 1994; McCarthy, Nobre, Bentin, & Spencer, 1995; Nobre & McCarthy, 1995). What imaging data seem to agree on is that a widely distributed network, concentrated in the left perisylvian region, is involved in reading.

Surprisingly few imaging studies have concentrated on reading in dyslexic subjects. Rumsey et al. (1987) reported that dyslexic subjects, reading and categorizing isolated words, showed higher regional cerebral blood flow (rCBF) in the left than right hemisphere, not paralleled in the control subjects. This asymmetry was suggested to reflect inefficient linguistic processing or inadequate bihemispheric integration. The better spatial resolution of positron emission tomography (PET) compared with rCBF has been used in at least two studies. Reading aloud isolated words at slow presentation rates was shown to produce higher bilateral blood flow in the inferior occipital cortex in dyslexic subjects compared to control adults, possibly reflecting inefficient wordform processing in dyslexics (Gross-Glenn et al., 1991). Further, dyslexic subjects displayed a tendency toward greater right than left inferior occipital flow, whereas control subjects exhibited the opposite trend. In a recent PET study, Rumsey et al. (1997) reported less activity in the bilateral temporal cortices in dyslexic adults than in normal subjects while reading aloud isolated words and nonwords and interpreted their findings as a dysfunction of temporal areas. Thus, hemodynamic studies have suggested variable patterns of differences in cerebral involvement between dyslexic and normal readers.

In addition to the spatial distribution, the temporal pattern of activation may also differ between dyslexic and control subjects while they read. The hemodynamic techniques lack the temporal resolution to follow the progression of activation from one brain area to another. Electroencephalography (EEG) and MEG allow noninvasive imaging of brain activation with millisecond time resolution. EEG measures the electric potential and MEG measures the magnetic field associated with synchronous activation of thousands of cortical pyramidal cells (Hämäläinen, Hari, Ilmoniemi, Knuutiila, & Lounasmaa, 1993). The skull and the scalp distort the electric potential but not the magnetic field, and therefore, localization of the active cortex is more straightforward with MEG than with EEG. The combination of accurate temporal and good spatial resolution makes MEG an ideal tool for studying developmental disorders in which differences in cortical function between the impaired and normal subjects may be quite subtle.

The only MEG study thus far conducted to investigate dyslexic readers used isolated words (Salmelin et al., 1996). Differences between dyslexic and control adults were found in the activation of three brain areas. The inferior occipito-temporal border in the left hemisphere
was active between 100 and 200 msec after word presentation in control subjects but not in dyslexic readers. Based on previous intracranial studies (Nobre, Allison, & McCarthy, 1994), the authors concluded that the missing or delayed response in dyslexic subjects might reflect impaired visual word recognition. Also, the activation in the left temporal area between 200 and 600 msec after word presentation seen in normal readers was mostly lacking in reading-disabled subjects. This could indicate inefficient activation of the phonological form of the word in dyslexic readers. The third difference was discovered in Broca’s area where an unexpectedly early activation was discovered in dyslexic subjects. This response was suggested to reflect a compensatory top-down process in which the dyslexic adults were passing the defective visual and phonological word form perception by trying to guess the identity of the words. A remarkably similar activation pattern (i.e., relative under-activation in left posterior regions and overactivation in left anterior regions) was recently reported in dyslexic readers using functional magnetic resonance imaging (fMRI) (Shaywitz et al., 1998).

**Detecting MEG Responses Associated with Semantic Processing**

The central problem in studies comparing normal and dyslexic subjects is our limited understanding of the cortical events involved in higher cognitive functions in general and reading in particular. Based on previous lesion and functional mapping studies, we can only make assumptions of what the differences between control and dyslexic subjects actually reflect. Because even the smallest changes in stimulus parameters may change brain activation dramatically (Price et al., 1994), comparisons are further complicated by differences in experimental setups between studies. Thus, the measurement itself should ideally be able to indicate the functional role of the cortical areas displaying differences between groups.

Event-related potential (ERP) studies have revealed a prominent negative brain potential about 400 msec after word presentation, which increases in amplitude with the amount of unexpected lexical/semantic information a word (or a picture) contains (Osterhout & Holcomb, 1995). In sentence contexts, the N400 response is largest to semantically inappropriate sentence-ending words (Kutas & Hillyard, 1980), smaller to semantically appropriate but unexpected words, and the least negative to words that are highly expected in the sentence context (Connolly, Phillips, & Forbes, 1995; Kutas & Hillyard, 1984). Dyslexic and control subjects have been compared during reading for size and peak latency of the N400 response: Brandeis, Vitacco, and Steinhausen (1994), using sentences that ended with an expected or an inappropriate final word, discovered that the peak latencies of the global field power curves for the inappropriate endings 400 to 600 msec after word onset were delayed in dyslexic compared with control children. No differences between the two groups were found in signal strength or topography. However, because the poor spatial resolution of EEG does not readily support anatomical localization and, thus, accurate description of the response, the spatio-temporal pattern of activation elicited in dyslexic individuals by reading comprehension has remained inconclusive. A recent fMRI study by Shaywitz et al. (1998) used carefully selected tasks that would allow the critical subtractions for the detection of lexical/semantic activation in dyslexic readers. However, this aspect of the dataset was not presented in detail.

A recent MEG study on normal adults (Helenius et al., 1998) identified consistent cortical generators for the N400 response. This study used four types of sentence-ending words, graded by their appropriateness to the preceding sentence context: Sentences presented visually one word at a time ended with (1) the expected most probable word (2) a word that was semantically appropriate but unexpected (rare) (3) a semantically inappropriate word that began with the same letters as the expected ending (phonological), or (4) with a word both orthographically and semantically inappropriate to the preceding sentence context (anomalous) (Connolly et al., 1995).

Cortical activation that differentiated between the semantically congruous and incongruous sentence-ending words was in most subjects localized in the left hemisphere, particularly in the left superior temporal cortex. This activation was equally pronounced and long lasting in response to the anomalous and phonological semantically incongruous sentence-ending words, weaker and more short-lived to rare unexpected words, and totally missing to the probable expected endings.

In the present study, the paradigm used by Helenius et al. (1998) was employed to determine timing, spatial distribution, and strength of cortical activations reflecting semantic congruity of the sentence-ending word in dyslexic subjects. We wanted to establish (1) whether the activation patterns differ between dyslexic and fluent readers and (2) how brain activations of dyslexic subjects in this naturally reading situation relate to findings of the previous MEG study (Salmelin et al., 1996) where isolated words were used as stimuli.

**RESULTS**

Figure 1 depicts 122-channel magnetic and 3-channel electric responses in one control subject evoked by reading the four types of sentence-ending words. Prominent variation of magnetic field was detected over several brain areas, but clear differences between the four sentence conditions were found especially over the left temporal area. The MEG response seen in the left tem-
poral sensors was influenced by the semantic appropriateness of the final word to the overall sentence context: Activation was strongest to the two semantically inappropriate sentence endings (anomalous and phonological), clearly weaker to rare endings, and totally missing to the probable endings. Activation in the left temporal region peaked around 400 msec (N400m) after the presentation of the semantically inappropriate sentence-ending words. Within the same time range, a negative-going N400 deflection following the unexpected sentence endings was also obvious in the three EEG channels.

Localization of N400m Response in Control and Dyslexic Readers

The 122-channel MEG responses were decomposed into the time behavior of activation in distinct brain areas by modeling the active neuronal populations as equivalent current dipoles. From the whole-head data of Figure 1, several reliably activated brain areas could be identified in this control subject (Figure 2) starting in the occipital lobe and later involving also the temporo-parietal areas. When the activation strengths in the distinct brain areas were followed as a function of time for the four types of sentence-ending words, only the sources in the left temporal (two sources) and right temporal areas (one source) showed significant differences (at least twice the size of the prestimulus noise level) between anomalous (and phonological) and probable sentence endings. In this control subject, the semantically sensitive activation in the left hemisphere began about 50 msec and peaked about 25 msec earlier than the activation in the right hemisphere.

Sources 6 and 7 in the left temporal area displayed similar time behavior, which increases the localization uncertainty of these sources; the 95% confidence limit for localization of source 7 (anterior-posterior 4 mm, inferior-superior 9 mm) does not exceed the size of the dots in Figure 2. The orientation of current flow is a useful additional indicator in identifying the active brain areas. The currents detected by MEG are produced mainly in the fissural cortex (Hämäläinen et al., 1993).
Thus, source 7 in Figure 2, with the orientation of current flow perpendicular to the course of the sylvian fissure, reflects activation in the upper portion of the superior temporal gyrus (rather than, say, in the lower portion of the central sulcus).

Figure 3 illustrates the N400m sources in all control and dyslexic subjects. Semantically sensitive source areas were found to be distributed over a wide area especially in the left but also in the right perisylvian cortex. In both groups, each subject had on average two N400m sources. The brain area that most consistently displayed semantic sensitivity was in the middle portion of the left superior temporal cortex: Seven of the ten control subjects tested and seven of the eight dyslexic participants had a source in this area. In two control subjects left-hemisphere N400m sources were localized in the frontal rather than superior temporal cortex. Five control and two dyslexic subjects had also additional sources more posteriorly either in the left parietal or temporal areas. Only one control subject had a reliably localizable source exclusively in the right hemisphere, and one dyslexic subject did not have any source that would have displayed reliable differences between anomalous and probable sentence endings. It would appear that the dyslexic individuals had more dispersed distribution of activation in the right hemisphere and more concentrated and scarce distribution in the left hemisphere than the normally reading subjects. However, more subjects would be needed to determine the reliability of this difference.

In control subjects more N400m sources were found in the left than in the right hemisphere (15 versus 5; $\chi^2(1) = 5.0, p < 0.03$). In dyslexic subjects, the difference between the left and right hemisphere failed to reach statistical significance, although the absolute number of N400m sources in the left hemisphere ($n = 9$) was more than twice the number of sources in the right hemisphere ($n = 4$). Accordingly, the distribution of N400m sources between the hemispheres did not differ in the two subject groups (Fisher exact probability test).

As shown in Figure 2 for one control subject, the sources tended to peak earlier in the left than right hemisphere: In all of those four control and four dyslexic subjects who had an N400m source both in the left superior temporal cortex and in the right hemisphere, the left-hemisphere sources peaked earlier (414 versus 441 msec in the control and 491 versus 582 msec in the dyslexic subjects for the inappropriate sentence endings). Thus, although the small number of right-hemisphere observations prevents statistical comparisons between hemispheres, the data strongly point to a true difference between the time behavior of the left and right hemispheres. Furthermore, because the latency difference between hemispheres seemed to be larger in dyslexic than in control subjects, the role of the right hemisphere activation may be different in the two groups.

The average depth of the sources in the left superior temporal area was 16 mm below the outer surface of the

Figure 2. Location and time behavior of the brain areas (numbers 1 to 8) activated by sentence-ending words in one control subject. The orientation of current flow within each localized source area is shown with a small tail. The curve at the bottom is the goodness-of-fit value, which indicates the percentage of the magnetic field variance explained by the present model. The line types are the same as in Figure 1.

Figure 3. Brain areas sensitive to the semantic appropriateness of the sentence final word in all control and dyslexic subjects.
brain in control and 12 mm in dyslexic subjects, with no significant differences between groups. The distance between N400m sources and areas activated by tones was on average 24 mm for control and 11 mm for dyslexic subjects. In dyslexic subjects, the mean error of localization (the 95% confidence limit) of the semantically sensitive sources was larger than in control subjects (7 × 7 × 7 mm³ versus 5 × 5 × 5 mm³; t(12) = −2.5, p < 0.03).

In-Depth Analysis of Activation in the Left Superior Temporal Cortex

Figure 4b illustrates the time behavior of the N400m sources localized in the left superior temporal cortex (Figure 4a) in one control and in one dyslexic subject. Compared with the control subject, the semantically sensitive activation detected in the dyslexic subject was clearly weaker and began and peaked later. Also, the dyslexic subject showed slightly weaker activation to the phonological endings than to the anomalous endings, unlike the control subject, who exhibited equally strong activation to the two types of semantically incongruous sentence endings.

The individual waveforms were scaled with respect to the peak amplitude in the phonological condition and averaged across those control and dyslexic subjects who had a source in the left superior temporal area (Figure 4a); in one control subject who had two N400m sources in this area, only the one that peaked earlier was included in the average. The mean responses of dyslexic subjects were then further scaled with respect to the average absolute activation strength in the control subjects. Because the activation to probable endings did not differ significantly from the prestimulus noise level in either of the groups (mean ± SEM 5.4 ± 1.5 nAm in control subjects and 4.4 ± 1.5 nAm in dyslexic subjects), the probable endings were not included in the further analysis of differences between sentence types.

Figure 4c shows the mean scaled responses of the dyslexic subjects to the three kinds of unexpected sentence endings, with the mean responses of the seven control subjects depicted with dotted lines in the background. A statistically significant main effect of Subject group on activation strengths was discovered; activation was stronger in the control than dyslexic subjects (F(1, 12) = 8.2, p < 0.01). Planned contrasts revealed that the difference between groups reached significance for phonological endings (F(1, 12) = 12.4, p < 0.004; 20.7 ± 2.4 nAm for control and 10.6 ± 1.6 nAm for dyslexic subjects) and for rare endings (F(1, 12) = 8.0, p < 0.02; 14.2 ± 2.0 nAm for control and 6.7 ± 1.7 nAm for dyslexic subjects). For anomalous sentence endings, the difference approached significance (F(1, 12) = 4.1, p < 0.07; 20.5 ± 2.4 nAm for control subjects and 13.7 ± 2.3 nAm for dyslexic subjects).

The main effect of Sentence type was significant both
when analyzed over all subjects \((F(2, 24) = 43.8, p < 0.0001)\) and when the average activation strength for the three types of unexpected sentence endings were analyzed separately for both groups \((F(2, 12) = 23.4, p < 0.001)\). The rare endings elicited a significantly smaller response than either anomalous (in dyslexic subjects, \(p < 0.001\)) or phonological endings (in dyslexic subjects, \(p < 0.0002\)). However, in control subjects the responses to the two types of semantically inappropriate endings did not differ (Helenius et al., 1998), whereas in dyslexic subjects the planned contrast revealed that the response to phonological endings was significantly smaller than to anomalous endings \((p < 0.04)\). Accordingly, the Sentence type by Subject group interaction approached significance \((F(2, 24) = 2.9, p < 0.08)\).

A statistically significant main effect of Subject group on the time behavior of the activation was discovered; the left superior temporal activation was significantly delayed in dyslexic subjects with respect to control subjects \((F(1, 12) = 19.3, p < 0.0009)\). This difference remained highly significant when the ascending (from 0 to 100% of the peak amplitude; \(F(1, 12) = 15.5, p < 0.002\)) and descending (from 100 to 0% of the peak amplitude; \(F(1, 12) = 14.4, p < 0.005\)) slopes were tested separately. Planned contrasts at different time points \((0\%\)–\(100\%\) on both sides of the maximum amplitude) revealed that the responses of the dyslexic subjects were significantly delayed at every measurement point. Even the 0% time point, the onset of semantically sensitive activation, differentiated between control and dyslexic subjects despite the large variance in this measure \((F(1, 12) = 4.91, p < 0.05)\).

In both groups the response latencies were significantly different for the three types of sentence endings, both when the overall time behavior of the sentence types was compared \((F(2, 12) = 20.6, p < 0.0005\) in control subjects and \(F(2, 12) = 8.4, p < 0.01\) in dyslexic subjects) and when the ascending \((F(2, 12) = 6.0, p < 0.02\) in control subjects and \(F(2, 12) = 5.7, p < 0.02\) in dyslexic subjects) and descending slopes \((F(2, 12) = 22.9, p < 0.0002\) in control subjects and \(F(2, 12) = 5.1, p < 0.03\) in dyslexic subjects) were tested separately. In control subjects, the ascending slope for phonological endings was significantly delayed compared with both anomalous \((p < 0.04)\) and rare endings \((p < 0.006)\), with the difference emerging when 75% of the peak activity had been reached. In dyslexic subjects the ascending slope of the phonological response was also delayed with respect to rare sentence endings \((F(1, 24) = 9.5, p < 0.02)\), but none of the separate contrasts analyzed at different time points reached significance.

In control subjects, the activation faded away (the descending slope of the activation) earlier for rare sentence endings than for anomalous endings, and the activation lasted about 30 msec longer for phonological endings than for anomalous endings (Helenius et al., 1998). In dyslexic subjects, only the difference between rare and phonological endings reached significance on the descending slope \((F(1, 24) = 8.8, p < 0.03)\). Thus, although both groups showed roughly the same time behavior for the three sentence types, as indicated by a nonsignificant Subject group by Sentence type interaction, many of the differences found in control subjects did not reach statistical significance in the dyslexic subjects due to greater variability within the group.

**Analysis of the ERPs in the N400 Time Window**

Figure 5 illustrates the grand average ERPs of all control and dyslexic subjects to the four different sentence types. Statistical analysis was performed on peak amplitudes measured between 300 and 600 msec after word onset from individual subjects. The main effect of Subject group on peak amplitudes failed to reach statistical significance irrespective of whether the analysis was done across all three electrodes or separately for each electrode. In control subjects, the response strengths to the four types of sentence-ending words varied significantly when analyzed over the three electrodes \((F(3, 24) = 7.4, p < 0.003)\). Planned contrasts revealed that the responses to both anomalous and phonological endings were significantly more negative than the response to probable endings \((p < 0.003)\). When the three electrodes were analyzed separately, the differences between semantically inappropriate and expected endings were significant only at Cz and Pz.

Consistent with the pattern described for control subjects, the response strength in the dyslexic subjects varied according to sentence type \((F(3, 15) = 12.1, p < 0.003)\). Dyslexic subjects showed significantly more negative responses to anomalous \((p < 0.01)\) and phonological endings \((p < 0.007)\) than to probable end-

![Figure 5. Averaged ERP responses of control and dyslexic subjects in electrodes Fz, Cz, and Pz. In one control and two dyslexic subjects the responses in Pz were not included in the averages or in the further analysis because of poor signal quality.](image-url)
ings. In addition, the responses to rare endings were significantly less negative than the responses to both anomalous and phonological endings (p < 0.01). But, unlike in control subjects, these differences were significant at all electrodes in the dyslexic subjects.

The statistical analysis of the latency of the N400 response was performed on three time points measured from individual subjects’ responses: the time point where the responses started to differ from the probable endings, the latency of the peak negativity, and the time point where the responses rejoined the probable endings. When control and dyslexic subjects were compared across all sentence types, electrodes, and time points, the main effect of Subject group remained nonsignificant, but a significant Sentence type by Subject group interaction emerged (F(2, 26) = 3.7, p < 0.05). When between-groups planned contrasts were performed separately for each electrode, a significant difference was found for electrode Pz: The latency of the peak amplitudes to anomalous (p < 0.03) and rare endings (p < 0.02) was delayed in dyslexic individuals and the response to the anomalous words also terminated later than in control subjects (p < 0.02).

In control subjects, the N400 response to the three unexpected sentence endings did not differ in the overall time behavior, whether tested over all three electrodes and three time points or when the electrodes and time points were inspected separately. In dyslexic subjects, the only statistically significant time difference between sentence types was discovered at the end point of the N400 response when tested over the three electrodes (F(2, 10) = 7.7, p < 0.02). Separate contrasts revealed that the response to rare endings terminated about 100 msec earlier than that to anomalous (p < 0.02) and phonological endings (p < 0.004).

To summarize, the ERPs, recorded with a limited set of electrodes, corroborated some of the MEG findings but failed to reveal many of the differences between groups evident in the MEG responses.

Analysis of the Early ERPs

Whereas only minor differences were observed in the N400 time range between the reading-disabled and normally reading subjects, the ERPs in the earlier time window between 150 and 350 msec were substantially different in the two groups (Figure 5): Unlike the control group that displayed a three-phased response complex P200-N240-P300 in the 150- to 300-msec range after word presentation, the dyslexic group seemed to totally lack the later negativity starting at around 200 msec and, instead, displayed a merged positivity in the 200- to 400-msec range. Only in two dyslexic subjects could negativity be seen around 250 msec in the individual ERP responses at the most posterior electrode Pz when the activity was averaged over all four sentence types; in the control group, this negativity was observable in half of the subjects. The strength of the positive deflection preceding the N400, integrated over the entire deflection from zero crossing to zero crossing and measured in individual subjects (found in eight control and in seven dyslexic subjects), was marginally stronger in the dyslexic than control subjects (t(13) = −2.1, p < 0.06).

Overall Spatio-Temporal MEG Pattern of Cortical Activation in Control and Dyslexic Subjects

Figure 6 displays all the identified source areas irrespective of their sensitivity to semantic appropriateness in the posterior brain regions in control and dyslexic subjects. The peak of the activation was determined as either the time point where the absolute maximum strength of the response during 0 to 900 msec after stimulus onset was reached or, if the source had two distinct peaks with the difference in strength less than 25%, as the latency of the first peak. No statistically significant differences were discovered in the number of control and dyslexic subjects having active sources in the left and right occipital and occipito-temporal areas (dashed rectangles) 100 to 200 msec after word onset. When the analysis was extended to 300 msec, fewer dyslexic subjects (n = 3, two sources in one dyslexic subject) than control subjects (n = 8) were found to show activation in the left occipito-temporal border. This difference did not, however, reach significance, unlike the differences in other posterior brain regions.

Figure 7 shows all the sources that were reliably localized in the left and right frontal, temporal, and parietal lobes (dashed lines) in control and dyslexic subjects. Some of the sources are the same as the semantically
sensitive sources shown in Figure 3. No statistically significant differences between groups were discovered in the number of subjects that had an activation peaking in the time windows 200 to 400 msec and 400 to 600 msec after final word onset in any of the three lobes in either hemisphere.

The Correlation Between MEG Signals and Behavioral Data

Figure 8 depicts the relationship between the behaviorally measured word-recognition speed and the cortical response to anomalous sentence endings in those control and dyslexic subjects who had an N400m source in the left superior temporal cortex (Figure 4a). The time point where the activation to anomalous endings had reached 50% of maximum amplitude was selected to represent the latency of activation. The correlation was significant in the control subjects (Pearson $r = 0.84, p < 0.02$) but not in the dyslexic subjects ($r = 0.50$, n.s.).

DISCUSSION

Localization of Semantic Activation in Dyslexic Versus Normal Readers

In this study the functional anatomy of reading in dyslexic individuals was investigated employing a paradigm that has recently been used to tap semantic activation in normal adult readers (Connolly et al., 1995; Helenius et al., 1998). The present study showed no differences in the spatial distribution of semantic activation between the reading-disabled and normally reading subjects: As seen in control subjects (Helenius et al., 1998), dyslexic adults displayed the highest absolute number of semantically sensitive sources in the left superior temporal cortex, with fewer sources identified in the left posterior parieto-temporal areas and in the right hemisphere. Thus in both groups, the middle portion of the left superior temporal cortex was most consistently activated by unexpected but not by expected sentence-ending words.

Developmentally dyslexic children have often been suggested to have an abnormal pattern of lateralization: According to Bakker (1992), reading problems in some dyslexic children are caused by an inability to shift from right-hemispheric guided to left-hemispheric guided reading strategies. Bryden (1988) reviewed more than 50 dichotic listening, visual half-field, and tactual dichhaptic studies and found some support for the association of reading disability with abnormal lateralization of language. In the present study, however, the hemispheric balance did not differ between dyslexic and normal readers when the semantic activation during reading was monitored simultaneously in the spatial and temporal domains. Naturally, negative findings should be regarded with caution due to the small number of subjects.

The Latency and Strength of Semantic Activation

Although no differences were discovered in the spatial distribution of semantic brain activation, clear differences in time behavior and activation strength were
evident between the subject groups. The activation sensitive to semantic appropriateness of the sentence final word was significantly delayed in the dyslexic readers compared with the control subjects—this was especially evident in the MEG data and was observed also as a delayed peak latency of N400 in our ERP data. This observation agrees with previous ERP studies on dyslexic (Brandeis et al., 1994) and on language-impaired children (Neville, Coffey, Holcomb, & Tallal, 1993), reporting a delayed peak of the N400 response while reading. In addition, the MEG results in the present study showed that also the onset of the N400 was delayed in dyslexic adults. This indicates that presemantic processing stages are effected in dyslexic readers, as also suggested by an earlier study (Salmelin et al., 1996).

A previous ERP study using a sentence reading task in dyslexic children found no difference in the strength of the N400 response (Brandeis et al., 1994), and this finding is supported by the present ERP recordings on dyslexic adults. However, in the MEG data the semantic activation was significantly weaker in the dyslexic than in the control subjects. An obvious reason for the inconsistency of ERP and MEG findings is the better spatial resolution of MEG. As the skull and the scalp distort the electric potential, source localization is rarely done using EEG: The electric N400 is a combination of signals from various brain areas irrespective of their sensitivity to semantic appropriateness. However, with MEG the activation in a particular area can be separated from other cortical (or subcortical) signals. Consequently, the sensitivity to even subtle differences is enhanced. The reduced strength of the MEG response suggests that the semantic cortical network involved in comprehension was either activated less synchronously or that a smaller number of neurons were engaged in the dyslexic than control adults. Thus, although the relationship between the strength of brain activation and efficacy of a behavior is by no means straightforward, the delayed and reduced activation in dyslexic readers suggests nonoptimal engagement of the conceptual neural network in reading.

### Differences of Reading Styles Between Control and Dyslexic Subjects

The results of Helenius et al. (1998) indicate that in fluent readers the phonological words (sharing the initial graphemes and phonemes with the most expected word) did not delay, to any great extent, the beginning of the semantically sensitive brain activation with respect to the other word types. Although the deceptive beginnings clearly influenced the time behavior of semantic activation, this could only be seen after the word-level semantic processing had been launched by lexical input or alternatively at the beginning of the contextual integration stage.

The onset of semantic activation for phonological final words was not delayed in the dyslexic subjects either. Such a delay, if it exists, could be masked by the large variance among dyslexic individuals in this small sample. However, the difference in strength between the anomalous and phonological endings does suggest that dyslexic subjects relied on sublexical word recognition: At least occasionally, the dyslexic adults mistook the correctly beginning word for the one they had expected, and consequently, the strength of the response to phonological endings was reduced with respect to anomalous endings. Another explanation for the reduced strength could be that within each dyslexic individual the latency of the semantic activation varied greatly for the phonological sentence-ending words, and thus, in the averaged responses the peak of activation was flattened. Based on the present MEG data it is not possible to decide whether phonological words were occasionally misread or whether the speed with which the words were recognized varied more for phonological than, for example, for anomalous words within dyslexic individuals. However, using behavioral methods, this question should be possible to settle.

Reading acquisition has been suggested to proceed through three stages (Frith, 1985). At the first stage, a child recognizes a limited number of words based on their visually salient features. At this stage the order of features is not yet emphasized. At the second stage, the child starts to use alphabetic rules and recognizes words by converting graphemes to phonemes from left to right. At the last stage, word recognition is again dominated by visual recognition of wholes, as polymorphemic words are recognized in syllabic or morphemic units. Reading processes in adult readers are often interpreted in terms of the dual-route model (Coltheart, 1978; Coltheart et al., 1993): Unfamiliar words and nonwords not stored in the orthographic lexicon are thought to be recognized by converting the word, grapheme by grapheme, into the corresponding phonemes (using the phonological route). On the other hand, familiar words are stored in an orthographic lexicon, and they are recognized via lexical access by direct activation of the semantic system (using the lexical route), without laborious phonological conversion. However, this dichotomy may not be as clear as originally suggested (see the opposing evidence of, e.g., Kay & Marcel, 1981; Lukatela & Turvey, 1994a, 1994b; Tousman & Inhoff, 1992).

According to Frith (1985), developmental dyslexia in children is characterized by a failure in mastering the second stage in reading acquisition: Due to deficient phonological skills (Bradley & Bryant, 1983; Shaywitz, 1996; Wagner & Torgesen, 1987), an awareness of the individual phonemes constituting spoken words does not develop and thus, the grapheme-to-phoneme conversion becomes hard to learn and remains laborious even later on. Consequently, reading will not progress further into the last level. With respect to brain activation this might indicate that the brain areas specializing in visual word form analysis (Halgren et al., 1994; Nobre et al.,...
where words of approximately the same length were poral, and left inferior frontal areas. In the present study, statistically significant in the left occipito-temporal, left tem-

The differences were statistically significant in the left occipito-temporal, left tem-

Adults viewing isolated seven- and eight-letter words, with no supporting context. The differences were statistically significant in the right occipito-temporal, left tem-

Accordingly, N400m latency separated the two groups because they were much more delayed in the behavioral word recognition task with a manual response (lexical decision) than could be expected based solely on their cortical responses. A reasonable interpretation is that in these dyslexic subjects the postsemantic (decision, executive, motor) processes were disproportionately prolonged, in addition to the affected presemantic stages. The N400m response latency is a rather pure indicator of delays due to presemantic processing, whereas the reaction time measure reflects both pre- and postsemantic processing stages. Accordingly, N400m latency separated the two groups more clearly than the reaction times. Combining behavioral and functional imaging data can thus further enhance the identification of the affected subcomponents in developmental disorders.

Relationship Between Behavioral and Cortical Measures

In the control group, the latency of the semantically sensitive activation in the left superior temporal cortex and behaviorally measured word recognition speed correlated statistically significantly. This can be expected assuming that both measures are closely related to the speed at which the meaning of the word can be accessed (Helenius et al., 1998; James, 1975). In the dyslexic group, correlation between cortical and behavioral latencies failed to reach significance. Especially two dyslexic subjects seemed to deviate from the pattern seen for control subjects because they were much more delayed in the behavioral word recognition task with a manual response (lexical decision) than could be expected based solely on their cortical responses. A reasonable interpretation is that in these dyslexic subjects the postsemantic (decision, executive, motor) processes were disproportionately prolonged, in addition to the affected presemantic stages. The N400m response latency is a rather pure indicator of delays due to presemantic processing, whereas the reaction time measure reflects both pre- and postsemantic processing stages. Accordingly, N400m latency separated the two groups more clearly than the reaction times. Combining behavioral and functional imaging data can thus further enhance the identification of the affected subcomponents in developmental disorders.

Reading Sentences Versus Isolated Words

In an earlier MEG study (Salmelin et al., 1996), clear differences in the spatio-temporal pattern of brain activation were discovered between control and dyslexic adults viewing isolated seven- and eight-letter words, with no supporting context. The differences were statistically significant in the left occipito-temporal, left tem-

To summarize, we were able to identify clear differ-

Helenius et al.  545
ences between developmentally dyslexic individuals and normal adults in the pattern of activation elicited by reading comprehension. Our findings stress the importance of using measurement techniques with good temporal and spatial resolution when studying cognitive processing in developmental disorders.

METHODS

Subjects

We studied eight developmentally dyslexic adults (three females; age 19 to 37 years, mean 27.6 years). Ten normal readers (five females; age 20 to 37 years, mean 27.0), whose results have been published earlier (Helenius et al., 1998), were used as control group. The two groups were comparable for age and educational background (Table 1). All dyslexic and normal adults were native speakers of Finnish. One of the dyslexic subjects was left-handed.

Although all the dyslexic subjects had overcome their initial difficulties in learning to read and write, they still experienced at least mild discomfort while reading. Their oral reading speed, word recognition speed, and performance in three other reading-related behavioral tests consistently differentiated them from the control group \((F(1, 16) = 9.6, p < 0.007, \text{Table 1})\).

Stimuli in the MEG Recording

The Finnish sentences used as visual stimuli in the MEG measurement were especially constructed to have high-cloze-probability sentence-ending words. Cloze probability (Taylor, 1953) was measured by asking 30 to 63 university students to fill in the most obvious ending word to each sentence. We used four different sentence types. In the \textit{probable} condition the sentence ended with the most probable final word (e.g., The piano was out of . . . tune). This ending was replaced by an unexpected, low-cloze probability, but semantically appropriate ending in \textit{rare} sentences (e.g., When the power went out the house became . . . quiet). \textit{Dark} would have been the most likely ending for this sentence. In \textit{anomalous} sentences the most probable ending was replaced by a final word that is totally inappropriate to the overall sentence context (e.g., The pizza was too hot to . . . sing). Sentences could also end with a word that was totally inappropriate to the overall sentence context but started with the same first few letters (two to four) as the most \textit{probable} ending (e.g., The gambler had a streak of bad . . . luggage). This sentence type was called \textit{phonological} because, due to the nature of the Finnish language (practically one-to-one correspondence between letters and sounds), the beginning of the word was both orthographically and phonologically the same as the most probable ending.

There were 100 sentences of each type. Each of the 400 sentences was presented only once to maximally engage the attention of the subjects. The sentences were composed of 4 to 10 words (mean 6.6, \textit{SD} 1.1). The final-word length was 5 to 13 characters (mean 7.7, \textit{SD} 1.8). Neither the mean length of the sentence nor the final-word length differed between the four sentence types. The mean cloze probability of the most often suggested ending was also equal (0.73 to 1.00) for all four sentence categories. Presentation of different categories was randomized, ensuring that the same sentence type was not presented more than three times in a row.

Stimuli in the Behavioral Tests

All subjects were tested in a variety of reading-related behavioral tests to detect residual signs of dyslexia. Oral reading speed was measured using a narrative printed on a sheet of paper; speed is reported as the number of

| Table 1. Background of Control and Dyslexic Subjects and Their Behavioral Profiles in Reading-Related Tests |
|-----------------|-----------------|-----------------|
|                  | \textit{Control Subjects} | \textit{Dyslexic Subjects} | \textit{Significance (univariate F test)} |
| Age (years)      | 27.0 ± 4.9       | 27.6 ± 6.9       | \textit{n.s.} |
| Educational level (years)a | 14.9 ± 3.3 | 13.1 ± 2.2 | \textit{n.s.} |
| Oral reading (words/min)b | 162.1 ± 26.0 | 94.9 ± 16.1 | \textit{p < 0.001} |
| Naming (item/msec)c | 474.2 ± 97.0 | 671.3 ± 138.7 | \textit{p < 0.003} |
| Word recognition (msec)d | 542.3 ± 62.1 | 846.7 ± 259.2 | \textit{p < 0.002} |
| Digit span (forward)e | 6.2 ± 1.3 | 5.8 ± 1.6 | \textit{n.s.} |
| Digit span (backward)e | 6.0 ± 1.6 | 4.5 ± 0.9 | \textit{p < 0.03} |
| \textit{a} Highest level completed. |
| \textit{b} The speed of reading aloud a narrative. |
| \textit{c} The speed of naming colors, digits, and letters presented in a matrix. |
| \textit{d} The average time to correctly identify Finnish words from orthographically legal pseudowords in a computerized task. |
| \textit{e} The number of digits correctly recalled. |
words read during 1 min. Word-recognition speed was tested with a computerized lexical decision task. The target words (Finnish word or an orthographically legal pseudoword) were preceded by a short (200 msec) prime that was either semantically related or unrelated to the target word. Half of the target words were presented against a partly masking background, but only the nondegraded (semantically or nonsemantically primed), correctly identified real Finnish words were used to estimate the word-recognition speed (number of averages 45 to 48 in both control and dyslexic subjects).

Naming speed, reflecting the ease with which the names of visually presented items are produced, has been shown to be reduced in dyslexic children (Denckla & Rudel, 1976; Wagner & Torgesen, 1987; Wolf, 1986; Wolf & Obregon, 1992) and in adolescents with a history of developmental dyslexia (Korhonen, 1995). Naming speed was measured using a 5 × 10 matrix consisting of numbers, letters, and colored squares.

The capacity to maintain verbal information in a short-term store while simultaneously dividing attention for processing incoming information is necessary for learning to read and write. The capacity of working memory has been occasionally shown to be impaired in dyslexic subjects (Jorm, 1983). The digit span backward and forward were measured using the standard Wechsler adult intelligence scale (WAIS) procedure (Wechsler, 1955).

MEG Recording

In the MEG recording the stimulus sentences were presented visually one word at a time for 330 msec with a 750-msec blank period between the words. The measurement room was dimly lit and the words, yellow on a black background, were projected onto a board in front of the subjects. The size of the word was on average 5.3 cm (visual angle 6°). Before starting it was ensured that all dyslexic subjects were able to read the sentences at the rate they were presented. Subjects were instructed to concentrate on the meaning of the sentences, to move their heads as little as possible, and to avoid excess blinking during the measurement. The measurement was conducted in six blocks, each lasting about 10 min. The blocks were interleaved with short (2 min) breaks. After three blocks, a longer 10-min break was necessary to save the measured raw data.

Subjects were tested in a magnetically shielded room with a Neuromag-122 whole-head device (Ahonen et al., 1993). EEG and electrooculogram (EOG) signals were recorded simultaneously with the MEG measurement. EEG recordings were made from Ag/AgCl electrodes over three midline sites at Fz, Cz, and Pz according to the International 10-20 System locations. Electrodes were referred to the left mastoid. Eye movements were monitored with horizontal and vertical EOG. Magnetic signals were bandpass filtered at 0.03 to 100 Hz, electric signals, at 0.1 to 100 Hz, and both digitized at 0.4 kHz. Signals were averaged on-line with respect to word presentation, separately for each sentence-ending word type, between −200 and 1000 msec in two sets (first three and last three measurement blocks). In the later off-line analysis, these two sets were further averaged. An epoch was discarded from the average when signals in one of the EOG channels exceeded 150 µV. To align brain activation with anatomical data, three head position indicator coils were attached to the subject’s head, behind both ears, and to the forehead. The locations of these coils with respect to three anatomical landmarks that can be easily detected in magnetic resonance (MR) images (nasion and preauricular points) were measured with a 3-D digitizer (Isotrak 3S1002, Polhemus Navigator Sciences, Colchester, VT, USA). Before the beginning of the first and fourth block the positions of the indicator coils with respect to the measurement helmet were determined from the magnetic signals produced by the coils.

Analysis of the MEG Data

Neuromag-122 applies planar gradiometers that detect the maximum magnetic signal just above the activated brain area. This feature helps the on-line interpretation of the measured signals because activation of different brain areas can already be, at least cursorily, followed from the averaged 122-channel responses (e.g., Figure 1). However, to investigate the measured signals more precisely, the magnetic field patterns were modeled with equivalent current dipoles (ECD) using a least-squares fit. ECD is an estimation of the center of gravity of the activated brain area, the strength of current flow in this area, and the direction of the current flow. As an estimation of the shape of the conducting volume (brain), necessary for localizing ECDs, a sphere model was used. The sphere model was created for each subject from his or her own MRI, emphasizing accuracy of the model in the temporal areas. For those three control and five dyslexic subjects who did not have an MRI available, an average male or female sphere model was used.

Because the activation patterns were highly overlapping for all three unexpected sentence types, the ECDs were identified in the condition and at the time point where activation of near-by cortical areas interfered as little as possible with the localization of activation. In a few cases, temporally and spatially highly overlapping activations were evaluated by first extracting a certain strong field pattern to reveal the underlying activity (Uusitalo & Ilmoniemi, 1997). Because the lengthy measurement time might have caused minor changes in the position of a subject’s head, the dipole modeling was done using the averaged data of the first (or the last) of the three measurement blocks, during which about half of the sentences were presented.

After localizing isolated dipoles (seven to nine in each
subject) using a subset of 12 to 20 channels (the channels in which the activation in question was clearly visible), all the dipoles were introduced into a multi-dipole model, keeping their locations and orientations fixed while their strengths were allowed to vary to achieve optimal explanation of the signals measured by the 122 sensors. Only those dipoles that had a maximum strength above 5 nAm and were at least twice the amplitude of the prestimulus noise level were accepted for the final model.

A source was considered to display semantic sensitivity (N400m source) if a difference twice the amplitude of the average prestimulus noise level was detected between probable and anomalous (or phonological) final words. For one source of one dyslexic subject this criterion was not quite met, but because the maximum strength of the activation to anomalous sentence-ending words differed from the prestimulus noise level, and that to probable words did not, the source was considered to display semantic sensitivity. To further characterize the time behavior of the left superior temporal sources the maximum amplitude and the time points when the source had reached 0, 25, 50, 75, and 100% of its maximum strength on both sides of the peak were measured in each individual.

After sources had been localized, they were combined with the anatomical MRIs. For those subjects who did not have MRI we used one female and one male brain on which the sources were projected. To identify the hand sensorimotor and auditory cortices of individual subjects, we also recorded evoked responses using self-paced right and left index finger lifts, performed about every 3 sec, and 50-msec 1-kHz tones, delivered alternately to the left and right ears every 1.2 sec.

Analysis of the ERP Data

In the EEG data, the N400 response was judged to be the most negative peak between 300 and 600 msec after word onset. This peak was identified in individual subjects’ ERPs at each electrode to the four types of sentence-ending words. The duration of the N400 response was studied measuring the starting and ending points of the differences between probable and rare/anomalous/phonological endings surrounding the peak negativity. If the waveforms between probable and unexpected endings started to diverge or did not rejoin before 0/1000 msec after word onset, these values were used instead of missing observations.

Statistical Analysis

The maximum strength of the N400m sources was tested separately in control and dyslexic groups using repeated measures analysis of variance (ANOVA) with the Sentence type as a within-subjects factor. The time behavior of the N400m sources was tested using repeated measures ANOVA, with Sentence type and Time point as the within-subjects factors. The Greenhouse-Geisser correction for sphericity violations was applied to the probability level where appropriate (epsilon < 1). Groups were compared using mixed-model ANOVAs with Sentence type or Sentence type and Time point as within-subjects factors and Subject group as a between-subjects factor.

The strength and the temporal pattern of the N400 responses were tested separately in control and dyslexic subjects using ANOVA (within-subjects factors Sentence type and Electrode or Sentence type, Time point, and Electrode) and applying the Greenhouse-Geisser correction to the probability level. Groups were compared using mixed-model ANOVA (within-subjects factors Sentence type and Electrode or Sentence type, Time point, and Electrode, and the between-subjects factor Subject group).

Acknowledgments

We are grateful to M. Illman for assistance during the MEG measurements and to R. Hari, K. Knuutka, and A. Tarkkainen for helpful comments on the manuscript. This work was supported by the Academy of Finland, Human Frontier Science Program, Natural Sciences & Engineering Research Council of Canada, Heart & Stroke Foundation of Nova Scotia, and March of Dimes Birth Defects Research Foundation. The MRIs were obtained at the Department of Radiology, Helsinki University Central Hospital.

Reprint requests should be sent to Päivi Helenius, Brain Research Unit, Low Temperature Laboratory, Helsinki University of Technology, PO Box 2200, 02015 HUT, Espoo, Finland, or via e-mail: paivi@neuro.hut.fi.

Note

1. P300 is not used here to refer to the P300 reported in classic odd-ball paradigms but only describes the waveform morphology.

REFERENCES


