Single word reading in developmental stutterers and fluent speakers

R. Salmelin, A. Schnitzler, F. Schmitz and H.-J. Freund

Brain Research Unit, Low Temperature Laboratory, Helsinki University of Technology, Espoo, Finland and Department of Neurology, Heinrich-Heine-University, Düsseldorf, Germany

Correspondence to: Riitta Salmelin, Brain Research Unit, Low Temperature Laboratory, Helsinki University of Technology, PO Box 2200, FIN-02015 HUT, Finland E-mail: riitta@neuro.hut.fi

Summary
Ten fluent speakers and nine developmental stutterers read isolated nouns aloud in a delayed reading paradigm. Cortical activation sequences were mapped with a whole-head magnetoencephalography system. The stutterers were mostly fluent in this task. Although the overt performance was essentially identical in the two groups, the cortical activation patterns showed clear differences, both in the evoked responses, time-locked to word presentation and mouth movement onset, and in task-related suppression of 20-Hz oscillations. Within the first 400 ms after seeing the word, processing in fluent speakers advanced from the left inferior frontal cortex (articulatory programming) to the left lateral central sulcus and dorsal premotor cortex (motor preparation). This sequence was reversed in the stutterers, who showed an early left motor cortex activation followed by a delayed left inferior frontal signal. Stutterers thus appeared to initiate motor programmes before preparation of the articulatory code. During speech production, the right motor/premotor cortex generated consistent evoked activation in fluent speakers but was silent in stutterers. On the other hand, suppression of motor cortical 20-Hz rhythm, reflecting task-related neuronal processing, occurred bilaterally in both groups. Moreover, the suppression was right-hemisphere dominant in stutterers, as opposed to left-hemisphere dominant in fluent speakers. Accordingly, the right frontal cortex of stutterers was highly active during speech production but did not generate synchronous time-locked responses. The speech-related 20-Hz suppression concentrated in the mouth area in fluent speakers, but was evident in both the hand and mouth areas in stutterers. These findings may reflect imprecise functional connectivity within the right frontal cortex and incomplete segregation between the adjacent hand and mouth motor representations in stutterers during speech production. A network including the left inferior frontal cortex and the right motor/premotor cortex, likely to be relevant in merging linguistic and affective prosody with articulation during fluent speech, thus appears to be partly dysfunctional in developmental stutterers.

Keywords: human; magnetoencephalography; language disorders; speech production; reading

Abbreviations: BA = Brodmann area; fMRI = functional MRI; MEG = magnetoencephalography; ROI = region of interest; SPECT = single photon emission computerized tomography; TOI = time window of interest; TSE = temporal spectral evolution

Introduction
Developmental stuttering is a sporadic disorder of speech production, which typically emerges at the age of 2–4 years. Stuttered speech is characterized by repetitions and prolongations of phonemes or syllables. For some individuals certain phonemes, especially consonants, may be particularly problematic. Dysfluency occurs most often at the beginning of a sentence or, more generally, when a new complete idea has to be expressed. Isolated words are stuttered less often. Stuttering occurs most frequently in self-initiated, self-paced discourse, but it is also evident when reading aloud. Increased emotional content in the discourse increases the frequency of stuttered events. The original incidence of stuttering is about 4%, but a vast majority of affected children show spontaneous recovery. About 1% of the population continues to suffer from severely stuttered speech even in adulthood, with a male-to-female ratio of 3 : 1 (cf. Starkweather, 1987; Bloodstein, 1995).

Theories of stuttering range from an isolated disorder of the speech motor system or auditory feedback to disrupted interaction in the multiple sensorimotor systems, and further
to stuttering as a manifestation of an unsuccessful alignment of subsequent sentence plans. Stuttering could be an outcome of periodic irregularities in the timing of muscle movements within the speech system (Zimmermann, 1980). When the background tension is high, as is often the case in stutterers (Freeman and Ushijima, 1978), the high-precision adjustments needed during speech become difficult to perform and the movements are jerky. Stutterers display poor coordination of antagonistic laryngeal muscles (Freeman and Ushijima, 1978), and they are systematically slower in initiating phonation than non-impaired speakers (Bloodstein, 1995). It has been suggested that specific neural correlates of the dysfunction of the motor system of stutterers can be found, e.g. in the coordination of speech movements by the supplementary motor area (Caruso, 1991). Stutterers also show abnormalities in rapid finger movements (Jäncke et al., 1995).

Malfunction of the auditory system during self-monitoring of speech has been suggested to underlie stuttering (Fairbanks, 1954). Normal speakers become dysfluent when exposed to delayed auditory feedback (Lee, 1951). In dichotic presentation of meaningful linguistic stimuli, a large proportion of stutterers fail to show the normally observed right-ear advantage (Curry and Gregory, 1969; Hall and Jerger, 1978). In fluent speakers, the left auditory cortex is more sensitive to the side of stimulation (right versus left ear), whereas the right auditory cortex is more sensitive in stutterers (Salmelin et al., 1998). Stutterers have also been reported to have difficulties in sound localization (Rousey et al., 1959). On the other hand, a possible central auditory processing deficit is likely to play only a minor role at the sentence-initial positions where stuttering most often occurs.

The small differences in motor or auditory function between fluent speakers and stutterers may reflect a more general interference between speech production and language formulation in stutterers (Starkweather, 1987). Karniol has recently proposed a framework for understanding stuttering, based on the idea that we produce complete sentences rather than single words when speaking or reading aloud (Karniol, 1995). Stuttering typically appears in children at the point where they work their way into more and more complex sentence structures (Starkweather, 1987; Bloodstein, 1995). Words are produced differently, in a shortened and modulated fashion, in a sentence than in isolation, and the word the word is produced depends on the sentence in which it is embedded. Sentences have suprasegmental features, including rhythm, melody and stress, that are largely determined prior to initiation of the utterance (see Karniol, 1995 for a review). The theory suggests that stuttering is caused by a misalignment at the border between subsequent suprasegmental plans. The more complex the utterance, the more difficult it is to superimpose a fundamental frequency contour, or prosody, on it.

Stuttering can be made to essentially disappear by reading in chorus with another person, or even by pacing the speech with a metronome (Johnson and Rosen, 1937). Both of these procedures reduce the need to build a prosodic contour for the expression, but the effect may also be due to the overall slowing down of speech production (Starkweather, 1987). Delayed auditory feedback (Lee, 1951; Soderberg, 1968) and auditory masking (Cherry and Sayers, 1956) can relieve stuttering. Again, the beneficial effect may arise from suppression of a defective auditory feedback system or from an overall slower speech rhythm resulting from the interference (Starkweather, 1987; Bloodstein, 1995). Stutterers can usually sing fluently, possibly because songs have no self-formulated propositional content (Starkweather, 1987).

Recently, the neural basis of developmental stuttering has been assessed using single photon emission computerized tomography (SPECT; Pool et al., 1991) and PET (Wu et al., 1995; Fox et al., 1996; Braun et al., 1997). Pool and colleagues reported global absolute blood flow reductions in stutterers as compared with fluent speakers in a resting condition (Pool et al., 1991). Braun and colleagues found significant differences between the stutterer and control groups when all subjects were speaking fluently (Braun et al., 1997). The cerebral function of fluent speakers and stutterers thus shows fundamental variance even in the absence of overt stuttering. To evoke dysfluent versus fluent speech, Wu et al. and Fox et al. asked the subjects to read aloud both self-paced and in chorus, while Braun et al. employed metronome pacing and recital of a familiar song in the fluent condition, and spontaneous narrative speech and sentence construction in the dysfluent condition (Wu et al., 1995; Fox et al., 1996; Braun et al., 1997). In these studies, the activation patterns differed between the subject groups during fluent speech, and between the fluent and dysfluent conditions in the stutterers, extending over a large number of cortical and subcortical areas. Results of previous neuroimaging studies seem to indicate that stuttering is associated with reduced activation of the left-hemisphere frontotemporal language areas (Wu et al., 1995; Fox et al., 1996; Braun et al., 1997). At the same time, the right hemispheric regions, including the motor and premotor cortices, show exceptionally strong blood flow in stutterers (Fox et al., 1996; Braun et al., 1997). Both PET (Fox et al., 1996) and magnetoencephalographic (MEG; Salmelin et al., 1998) studies have implied altered auditory cortical function in stutterers, particularly in the left hemisphere. Although the picture still remains rather unspecific with regard to the possible causes of stuttering, there is clear evidence for extensive functional differences in the brains of stutterers and fluent speakers.

We employed whole-head MEG to investigate the timing of cortical activation sequences in developmental stutterers and fluent speakers. As MEG combines excellent temporal resolution with good accuracy of localization of active cortical areas, it is a useful tool for characterizing the activation sequence from visual perception to oral output (cf. Salmelin et al., 1994) and for identifying cortical correlates of disorders in these processes. The subjects read aloud single words.
The vocalization was delayed by half a second to highlight perceptual and motor production aspects of the process, and to reduce the effect of mouth movement artefact on the early cortical activation patterns. This task is both behaviourally and experimentally straightforward and convenient. It was not expected that it would evoke much stuttering, except in the most severely affected individuals. The paradigm allows comparison of brain function in fluent speakers and stutterers during essentially identical overt performance. Earlier imaging results have shown differences in cerebral blood flow between stutterers and fluent speakers at rest and in fluency-evoking reading conditions (Pool et al., 1991; Braun et al., 1997). It is thus possible that the seeds for stuttering are present constantly but the threshold for overt dysfluency is exceeded only periodically (Bloodstein, 1995). Developmental stuttering is a particularly intriguing disorder as the deficit is only functional, without an obvious structural correlate. Distinct cortical activation patterns in stutterers and fluent speakers, associated with identical overt behaviour, could elucidate not only the neuronal basis of stuttering but also the steps necessary for normal speech production.

Methods

Subjects

Nine developmentally stuttering subjects (S1–S9; 22–53 years, mean 36 years, 7 males) and 10 fluent speakers (C1–C10; 25–52 years, mean 34 years, 8 males) gave their informed consent to participate in the study, which was approved by the ethical committee of The Medical Faculty of the Heinrich-Heine-University. All subjects were German-speaking and right handed as assessed by a handedness questionnaire (Annett, 1967) and a standardized hand performance test (Jäncke et al., 1995). Stuttering was very mild in one subject, mild in two, moderate in two, severe in two and very severe in two (Stuttering Severity Instrument; Riley, 1972). All our stutterers had tried at least one type of therapy for their speech problems at some point in their life, including conventional speech therapy (6 subjects), metronome-paced speaking (2 subjects), psychotherapy (3 subjects), relaxation techniques (1 subject) and fluency therapy (1 subject). No therapeutic interventions were administered while the subjects participated in this study.

All subjects were tested on the Wechsler adult intelligence scale (WAIS; Wechsler, 1991). The performance of both fluent speakers (115 ± 4.5, mean ± standard deviation) and stutterers (110 ± 7.2) was within normal range. There were no significant group differences.

Task

The stimuli were common German nouns, composed of 7–8 letters (48% concrete nouns, 42% abstract nouns, 10% with both a concrete and an abstract meaning). Each word was presented for 300 ms. After a blank interval of 500 ms, a question mark appeared for 2000 ms, prompting the subject to read the word aloud. The question mark was followed by a blank period of 2000 ms. The whole sequence (word–blank–question mark–blank) was thus repeated every 4.8 s. Altogether the stimulus set contained 250 different words.

To obtain functional landmarks in the auditory cortex and sensorimotor hand area, in separate runs, the subjects received 1 kHz, 50 ms tones every 1 s, alternately to the left and right ear, and performed self-paced index finger lifts approximately every 3 s, alternately with the left and right hand. Spontaneous brain activity during resting was recorded for 1 min when the subjects had their eyes closed and for 1 min when their eyes were open. The results of the auditory experiment have been reported separately (Salmelin et al., 1998).

MEG

Neuromagnetic signals reflect synchronous postsynaptic potentials in tens of thousands of pyramidal cells within a cortical patch on the order of a square centimetre. Because of the closely spherical symmetry of the human brain and skull, the detected MEG signals are mainly associated with electric current flowing parallel to the skull, i.e. with activation in the fissural cortex. A detailed description of the MEG method is given by Hämäläinen and colleagues (Hämäläinen et al., 1993).

The Neuromag-122™ whole-head MEG system (Neuromag, Helsinki, Finland) contains 122 sensors arranged on a helmet-shaped surface. Each sensor is composed of a pick-up coil, which collects the magnetic field associated with neuronal current flow, and a superconducting quantum interference device (SQUID). The SQUID transforms the magnetic field to voltage which can be measured with high accuracy (Ahonen et al., 1993). The planar gradiometers used in Neuromag-122™ detect the maximum signal immediately above an active cortical area.

Measurement procedure

The measurements were performed in a magnetically shielded room. The subject was seated on a chair, with the head supported against the helmet-shaped bottom part of the MEG apparatus. The words, white letters on a dark grey background, subtended a 4° visual angle on a back-projection screen placed at 1 m from the subject.

The MEG signals were recorded with a 0.03–130 Hz filter and digitized at 0.4 kHz. Both vertical and horizontal EOG (electro-oculogram) were recorded simultaneously. In addition, lip movements were monitored with EMG across the opposite corners of the mouth (orbicularis oris muscle); in two stutterers and in one control subject, technical problems prevented the collection of the lip EMG signals. The subject’s speech was registered with a microphone and stored on audiotape. The continuous MEG, EOG, EMG and microphone records were stored on magneto-optical disk for off-line analysis.
MEG signals were averaged on-line from -200 to +1500 ms with respect to stimulus onset. Epochs contaminated by eye or eyelid movements were rejected from the average. A minimum of 90–100 artefact-free epochs were collected for all subjects. The stimuli were presented in blocks of about 60 words, each lasting for 5 min. In fluent subjects, two blocks usually provided enough repetitions. In the two most severe stutterers, all four blocks had to be run to obtain the minimum number of non-stuttered and artefact-free epochs.

**Data analysis**

**Behavioural measures**

The verbal responses of the stutterers were evaluated after the measurement from the audiotape recording by two independent workers. A response was accepted as fluent when both referees unequivocally agreed on this classification. Stuttered responses had a clear repetition of a phoneme or syllable. If there was a hint of dysfluency but not full-blown stuttering, or if the two referees disagreed on the classification, the response was considered ambiguous. The ambiguous category never exceeded 20% of the trials. The subjects were also interviewed for their personal impression after the MEG recording.

Mouth movement and speech onset latencies were determined from burst onsets in the lip EMG and microphone records in all subjects. In each individual, the EMG and microphone signals were rectified and averaged with respect to word onset to obtain the overall shape of mouth muscle activity and vocalization. For calculating the mean shape across subjects, the individual signals were normalized by setting the maximum equal to 1.

**Time-locked evoked responses**

The MEG signals of the stutterers were re-averaged off-line for fluent, stuttered and ambiguous responses, from 200 ms before to 1800 ms after word presentation. The MEG signals were also averaged off-line with respect to lip movement (-1000 to +1000 ms) and microphone signal onsets (-1000 to +1000 ms) in all subjects. In the following text, these three different reference points will be referred to as word onset, mouth movement onset and speech onset. Before source analysis, the MEG data were further low-pass filtered at 40 Hz.

All subjects’ data were analysed individually. A muscle/tongue artefact coincided with microphone onset in seven of the 10 controls and in all stutterers. The conspicuous bilateral artefact pattern, which was at a maximum towards the rim of the helmet, was highly similar in all subjects. The artefact distribution was identified at the time point where there was the least evidence of simultaneous cortical responses, and the field pattern was removed from the MEG signals using the signal-space-projection method (Uusitalo and Ilmoniemi, 1997). Field patterns at other time intervals before and after the peak disturbance were visually inspected to verify that the actual cortical activation patterns were not affected by the signal-space-projection procedure. In two control subjects, the artefact could not be removed because of simultaneous strong cortical activity, but the cortical activation patterns could still be analysed satisfactorily. When source modelling was complete, the data were checked both in the original form and with the artefact removed.

The active cortical areas were modelled as current dipoles (Häimiläinen et al., 1993). The dipole’s location, orientation and amplitude represent the centre of gravity of the active cortical patch and the direction and mean strength of the current flow therein. The process of source modelling consists of continuous interplay between visual inspection of coherent local signal variations in the original responses, a search for clear dipolar field patterns in the analysis programme, and evaluation of how well the source model accounts for the measured signals (goodness-of-fit). The current dipoles were identified one by one, at time points where each specific field pattern was clearest. The sources were then brought into a multi-dipole model where the source locations and orientations were kept fixed while their amplitudes were allowed to vary as a function of time to best account for the signals measured by all the 122 sensors. The resulting time courses of activation in the cortical source areas are referred to as source waveforms, to separate them from the original 122 MEG sensor waveforms. The complete model included between nine and 13 sources in each individual, when responses were averaged with respect to stimulus onset, and between four and nine sources when the responses were averaged with respect to lip EMG onset. The goodness-of-fit varied between 70 and 90% across subjects and analysis intervals. Each source accounted for 18 ± 13% (mean ± standard deviation) in the goodness-of-fit value during the time interval when the source was most active, and a minimum of 5% within one hemisphere. The accuracy of source localization was on average 6 mm (95% confidence limit).

The individual source models were compared across subject groups. Based on clustering of the sources, the brain was divided into 12 regions of interest (ROIs): midline occipital cortex, left and right occipito-temporal, inferior frontal, superior temporal, inferior parietal and Rolandic cortex, and vertex (see detailed description in Results). The source waveforms were averaged across subjects within each ROI. When a subject had several sources in one region, the waveforms were added together. If a subject did not have a source within a certain ROI, the source waveform was set equal to zero, as the signal was apparently so small that no distinct source area could be identified. The waveforms in a specific ROI were included in further analysis only if at least half of either controls or stutterers, i.e. a minimum of five subjects, showed a response there. This criterion excluded the right superior temporal cortex from further consideration, leaving 11 ROIs for detailed analysis.
Within each ROI, time windows of interest (TOIs) were chosen as those time intervals where the averaged waveforms in fluent speakers and stutterers differed by at least their standard errors of mean. The source waveforms were analysed in 25 ms bins, corresponding to the smoothing effect of the 40-Hz low-pass filter. The time-locked responses are typically quite sharp (duration up to 100 ms) within the first 400 ms after stimulus onset and become temporally more widespread or sustained at longer latencies. Therefore, for an interval to qualify as a TOI, the mean source waveforms of the two groups were required to differ continuously for at least 50 ms (two adjacent 25 ms bins) at latencies 0–400 ms after stimulus onset. For latencies longer than 400 ms, the waveforms of the two groups were required to show a difference lasting for at least 100 ms. The requirement of a minimum difference of 100 ms was also applied for source waveforms averaged with respect to lip EMG onset. For each candidate ROI/TOI, the mean signal strength was calculated in each subject. Group differences in the mean source strengths were tested using one-way ANOVA (analysis of variance).

For estimating whether source strength differed from zero within a certain ROI/TOI, we used the base level of each waveform, i.e. the standard deviation within the prestimulus interval (–200 to 0 ms). The activation was taken to differ significantly from zero at \( P < 0.05 \), \( P < 0.01 \) and \( P < 0.001 \), when the amplitude exceeded 1.96, 2.58 and 3.29 SD, respectively. Again, to be accepted as a true response, a peak was required to be non-zero at least for 50 ms within the first 400 ms and for at least 100 ms at longer latencies.

**Cortical rhythmic activity**

In addition to the time-locked evoked responses, we analysed event-related modulation of cortical rhythmic activity. First, amplitude spectra were calculated for each subject in all measurement conditions (word reading, auditory stimulation, finger movements, resting) by advancing a 2.6 s window in 1.3 s steps through the entire non-averaged data set and averaging the resulting spectra. In all subjects, it was possible to identify four distinct spectral ranges (passbands): (i) 8–11 Hz (low 10 Hz); (ii) 11–15 Hz (high 10 Hz); (iii) 15–21 Hz (low 20 Hz); and (iv) 21–28 Hz (high 20 Hz), with the borders varying by 1–3 Hz across individuals. The event-related modulation of the cortical rhythms was analysed with the Temporal Spectral Evolution (TSE) approach (Salmelin and Hari, 1994b). The MEG signals were filtered through the four individually determined passbands described above, rectified (absolute value), and averaged with respect to stimulus onset. For quantification of the modulation, the TSEs were also calculated in the fixed passbands of 8–14 Hz (‘10 Hz’) and 16–28 Hz (‘20 Hz’) for all subjects.

Sources of the rhythmic activity were searched every 10 ms from 150 s of non-averaged signals (2 × 30 s during word reading and finger movements and 30 s while resting with eyes closed) filtered through the individually selected passbands, using subsets of sensors over the left and right hemispheres (covering the temporal lobe and the central sulcus), the posterior areas (posterior parietal and occipital cortex), and the vertex (Salmelin and Hari, 1994a). Rhythmic activity was concentrated close to the hand and mouth areas, and around the parieto-occipital sulcus and calcarine fissure. Additional clusters (1–2) in the posterior parietal cortex were seen in some subjects. During strong bursts of rhythmic activity, in particular, it was possible to model the generators as current dipoles. The hand area rhythms could be identified in any frequency range whereas in the mouth area the bursts concentrated in the 20-Hz range. Each distinct cluster (5–7 per subject) was represented by a single dipole. By forming a multidipole model, TSE curves of the cortical sources were obtained (Salmelin et al., 1995). Visual comparison of the whole-head TSE curves resulting from this dipole model with the original TSE curves showed that the selected generators explained the measured modulation of rhythmic activity in all subjects.

Statistical tests were performed on the source TSE waveforms in the hand and mouth areas (instead of original sensor outputs), using a mixed-model ANOVA (hemisphere × area × subject group).

**MEG and MRI**

Anatomical MRIs were available for all subjects. For presenting the functional MEG results on the MRIs, the two coordinate systems were aligned with the help of three small coils placed on the subject’s head prior to the measurement. Using a 3D digitizer (Isotrak 3S1002, Polhemus Navigation Sciences, Colchester, Vt., USA), the positions of these coils were determined with respect to three landmarks on the head, i.e. nasion and points just anterior to the ear canals, which are readily identified on the MRIs. The locations of the coils with respect to the MEG helmet were determined by briefly energizing the coils and calculating their locations from the magnetic field patterns.

The MRIs of all subjects were available both as slices and as surface renditions. Sources of the auditory and finger sensorimotor activations were superimposed on the subject’s MRIs. Their locations in the suprasylvian auditory cortex around Heschl’s gyrus and in the hand knob along the central sulcus (cf. Yousry et al., 1997) verified the correct alignment of the MEG and MRI coordinate systems. The sources of the word reading task were then superimposed on the individual MRIs. The sources were located at a depth of 12 ± 4 mm (mean ± standard deviation). For surface renditions, the sources were projected along the head radius, and adjusted to the correct sulcus with the help of the 3D MRI slices.

For comparison of active areas in the two subject groups, the sources were further transferred on a single subject’s brain. Care was taken to ensure that the source locations remained correct in relation to the sulcal structure and...
Fig. 1 Distribution of mouth movement onset latencies \((\text{upper panels})\) and speech onset latencies \((\text{lower panels})\), displayed as bar graphs, and the mean shape of mouth EMG \((\text{upper panels, black curve})\) and microphone signal \((\text{lower panels, grey curve})\) in \(\text{(A)}\) fluent speakers and \(\text{(B)}\) stutterers when they were fluent. The proportion per bar is given on the left vertical axis (left arrow) and the mean normalized signal strength (maximum set equal to 1 in each individual) on the right vertical axis (right arrow). The word was shown at 0 s (left vertical dashed line) and the question mark at 0.8 s (right vertical dashed line). The width of each bar is 0.1 s.

Results

Mouth movement and speech

Distribution of onset latencies and mean time behaviour of mouth muscle activity and speech signal are displayed in Fig. 1 across all fluent speakers (Fig. 1A) and for non-stuttered words in developmental stutterers (Fig. 1B). Mouth movement started on average 960 ms after word onset, i.e. 160 ms after the question mark prompt for vocalization, and overt speech began on average 290 ms later. Mouth movements lasted for \(\sim 1050\) ms and speech for \(\sim 650\) ms. Fluent speakers and fluently reading stutterers did not differ in any of these measures. Their overt performances were thus indistinguishable in this task.

Although muscle activity occasionally started before the question mark onset, both the maximum muscle activity and the microphone signal followed the vocalization prompt, verifying the correct performance of the task. Interestingly, while the speech onset latency was the same in fluent speakers and stutterers, the mouth movement onsets tended to spread into earlier latencies in stutterers than fluent speakers (bars in the upper row). Severity of stuttering did not correlate with mouth movement or speech onset or duration, or delay from mouth movement to speech onset.

As expected, the stutterers were mainly fluent in this task. Five stutterers were dysfluent at least occasionally (range 9–108 stuttered words), but only one of them had a large enough number of stuttered trials to provide an acceptable signal-to-noise ratio in the MEG signals. The ensuing analysis compares cortical activity of stutterers and control subjects when both were fluent.

Activity time-locked to word presentation

Figure 2 displays the whole-head MEG signals and the corresponding source waveforms in one fluent subject (C1) from 200 ms before to 1400 ms after word onset. The signals are strongest immediately above the active cortical area. Even from the whole-head data, one can readily recognize the early occipital and right fronto-parietal activations, and a persistent response over the left temporo-parietal cortex. A detailed source analysis (see Methods) revealed 11 reliable source areas which accounted for at least 80% of the MEG signal variance during most of the studied interval when all the 122 sensors were included. The medial and lateral occipital areas (sources 1, 2 and 3) were active first, followed by the right rolandic source 4, each for less than 100 ms. Movements lasted for \(\sim 1050\) ms and speech for \(\sim 650\) ms. The right (source 5) and left (source 7) inferior parietal and left superior temporal cortex (6) then started to participate, remaining active for 200–400 ms. A brief signal from the left inferior frontal cortex (source 8) was then followed by activation of the left posterior temporoparietal cortex (source 9). Activation of sources 10 and 11, reflecting involvement of the left posterior temporal cortex and the left dorsal premotor cortex, respectively, continued when the question mark appeared at 800 ms, prompting the subject to read the word aloud.

The original 122 MEG signals were resolved into the time behaviour of distinct cortical areas in all individuals. Figure 3 combines the source areas of all fluent subjects on the left and all stutterers (when they were fluent) on the right. The similarity of the sources was judged by the location and orientation of current flow. The distinct clusters of sources are indicated with different shapes and colouring of the markers.

The ROIs were identical in the left and right hemispheres. The ‘inferior frontal’ subregion (Fig. 3, white squares) included sources clustering predominantly in the frontal
Fig. 2 Left: MEG responses of fluent subject C1, averaged with respect to word onset (at 0 ms). The measurement helmet is viewed from above, flattened onto a plane, with the nose pointing upwards. Time is given in the horizontal axis (from 200 ms before to 1400 ms after word onset) and the variation of magnetic field in the vertical axis. Neuromag-122™ records maximum signal immediately above an active cortical area. The 122 sensors are arranged in 61 locations along the helmet. In each location, there are two overlapping sensors (see schematic heads in the upper right corner). The topmost one is most sensitive to ‘vertical’ currents flowing towards or from the vertex, and the lower one to ‘horizontal’ currents flowing around the perimeter. Right: Source analysis of the MEG signal distribution over the same time interval as on the left. Source strengths (in nanoamperemetre, nAm) are plotted as a function of time. The word onset at 0 ms and the appearance of the question mark prompt at 800 ms are indicated with solid vertical lines. The locations (dots) and orientations (tails) of the dipolar sources are shown on schematic drawings of the subject’s brain.

The operculum and insula, with posterior-to-anterior current flow (cf. Fig. 2, source 8, for single-subject data). This ROI was delimited from the inferior end by the superior temporal sulcus. The ‘superior temporal’ subregion (black triangles) included sources with current flow directed away from the vertex (cf. Fig. 2, source 6). The ‘frontoparietal’ subregion (Fig. 3, white circles) covered the primary somatosensory and motor cortices and the premotor and dorsal prefrontal cortex, delimited by the postcentral sulcus, precentral sulcus, and the inferior frontal sulcus. In the ‘frontoparietal’ ROI the current was mainly orientated along the posterior–anterior axis (cf. Fig. 2, sources 4 and 11). The ‘inferior parietal’ subregion (Fig. 3, white arrowheads) was defined by the postcentral sulcus, sylvian fissure, and the posterior ascent of the superior temporal sulcus, thus covering the angular gyrus and supramarginal gyrus. The current was typically orientated along the anterior-to-posterior axis (cf. Fig. 2, source 7). In the ‘occipitotemporal’ ROI (black circles) the current usually flowed towards the vertex (cf. Fig. 2, sources 2 and 3), and the area was confined by the ascending ramus of the superior temporal sulcus, the anterior occipital sulcus, and the lunate sulcus. The ‘occipital’ subregion (Fig. 3, black squares) included all parieto-occipital sources clustering close to the posterior midline. In the ‘vertex’ ROI (Fig. 3, white inverted triangles), the current flow was in the posterior–anterior direction and the sources concentrated frontally to the central sulcus. In addition, scattered sources were found in the left middle temporal cortex (Fig. 3, grey circles) but they did not form clear clusters, based on either location or orientation, or on the percentage of subjects showing the same type of response. Accordingly, further analysis was conducted in 11 ROIs: left and right inferior frontal, frontoparietal, inferior parietal, and occipitotemporal cortices, left superior temporal cortex, the occipital subregion, and over the vertex.

Figure 4 gives the mean ± standard error of the mean source waveforms in each of the 11 ROIs, averaged over fluent subjects and stutterers. The responses can be divided
Fig. 3 Source areas in fluent subjects (left) and stutterers (right), when the MEG signals were averaged with respect to word onset. The different shapes and colours of the symbols (white and black circles, triangles, squares and arrows) depict the grouping of sources into distinct ROIs (see text for details). The grey circles denote sources which do not belong to any well-defined cluster. The black diamond indicates the location of the hand sensorimotor cortex. The size of the symbols equals the mean accuracy of localization (6 mm, see Methods).

roughly into three stages. In the first stage, within 200 ms after word onset, the occipital (area 1) and left and right occipitotemporal cortices (areas 2 and 3) showed strong, transient signals, and also weaker responses after the question mark. In the second stage, at 200–600 ms, the persistent occipital activation was accompanied by responses in the left and right inferior frontal (areas 4 and 5), left superior temporal (area 6), and left and right inferior parietal subregions (areas 7 and 8). In the third stage, the left and right frontoparietal cortices (areas 9 and 10) and the vertex (area 11) became involved and remained active throughout the vocalization.

Fig. 4 Mean ± standard error of the mean (black curve and shading) source strengths as a function of time (cf. Fig. 2) in fluent subjects (left) and stutterers (right). The word and question mark onsets are indicated with solid vertical lines and the mouth movement (M) and speech (S) onsets with dashed lines. The black arrowheads denote the ROIs and TOIs (see text) where the responses of stutterers and fluent speakers differed significantly from each other. The studied ROIs are illustrated on the schematic drawings of brains on the left.
was active in seven of 10 control subjects but only in three stutterers. In controls (one source per subject). Note the involvement of in at least one of these ROIs, the peak latencies of the mean source strengths differed between stutterers compared to fluent speakers (Fig. 5C), with the centre of activity on average 1 cm inferior and anterior to the hand area, was involved in seven stutterers (not significant). The right frontoparietal source (Fig. 5C) showed activation both in the inferior frontal and frontoparietal ROIs. Comparison of the latencies of the left inferior frontal cortex (Fig. 5C) with the centre of activity on average 1 cm inferior and anterior to the hand area, was involved in seven control subjects (two sources in four subjects) but only in one stutterer (P < 0.02, Fisher’s exact test).

The sequence of activation in the left hemisphere was observed directly in the five fluent speakers and four stutterers who showed activation both in the inferior frontal and frontoparietal ROIs. Comparison of the latencies of the earliest peak activations in the two areas implied that the response sequence was from motor cortex to inferior frontal cortex in all four stutterers but from inferior frontal to motor cortex in four of the five fluent speakers (P < 0.05, Fisher’s exact test). In the nine fluent speakers who showed activation in at least one of these ROIs, the peak latencies (mean ± standard error of the mean) of the left inferior frontal and frontoparietal responses were 230 ± 27 ms and 348 ± 58 ms and in the eight stutterers 377 ± 44 ms and 197 ± 32 ms, respectively. A 2 × 2 mixed-model ANOVA resulted in a significant area-by-group interaction [F(1,7) =

<table>
<thead>
<tr>
<th>ROI*</th>
<th>TOI (ms)</th>
<th>Difference</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Occipital (1)</td>
<td>225–275</td>
<td>F &gt; S</td>
<td>n.s.</td>
</tr>
<tr>
<td>Occipital (1)</td>
<td>750–900</td>
<td>F &gt; S</td>
<td>n.s.</td>
</tr>
<tr>
<td>Left occipitotemporal (2)</td>
<td>775–950</td>
<td>S &gt; F</td>
<td>n.s.</td>
</tr>
<tr>
<td>Left occipitotemporal (2)</td>
<td>1100–1400</td>
<td>S &gt; F</td>
<td>P &lt; 0.05</td>
</tr>
<tr>
<td>Right occipitotemporal (3)</td>
<td>375–425</td>
<td>S &gt; F</td>
<td>n.s.</td>
</tr>
<tr>
<td>Right occipitotemporal (3)</td>
<td>1275–1400</td>
<td>S &gt; F</td>
<td>n.s.</td>
</tr>
<tr>
<td>Left inf. frontal (4)</td>
<td>200–275</td>
<td>F &gt; S</td>
<td>P &lt; 0.02</td>
</tr>
<tr>
<td>Right inf. frontal (5)</td>
<td>225–250</td>
<td>F &gt; S</td>
<td>n.s.</td>
</tr>
<tr>
<td>Right inf. frontal (5)</td>
<td>800–950</td>
<td>S &gt; F</td>
<td>n.s.</td>
</tr>
<tr>
<td>Left sup. temporal (6)</td>
<td>800–875</td>
<td>F &gt; S</td>
<td>n.s.</td>
</tr>
<tr>
<td>Left inf. parietal (7)</td>
<td>350–375</td>
<td>F &gt; S</td>
<td>n.s.</td>
</tr>
<tr>
<td>Right inf. parietal (8)</td>
<td>550–725</td>
<td>F &gt; S</td>
<td>n.s.</td>
</tr>
<tr>
<td>Right inf. parietal (8)</td>
<td>1175–1325</td>
<td>F &gt; S</td>
<td>n.s.</td>
</tr>
<tr>
<td>Left frontoparietal (9)</td>
<td>100–200</td>
<td>S &gt; F</td>
<td>P &lt; 0.03</td>
</tr>
<tr>
<td>Left frontoparietal (9)</td>
<td>400–475</td>
<td>S &gt; F</td>
<td>n.s.</td>
</tr>
<tr>
<td>Right frontoparietal (10)</td>
<td>325–550</td>
<td>F &gt; S</td>
<td>n.s.</td>
</tr>
<tr>
<td>Right frontoparietal (10)</td>
<td>1025–1400</td>
<td>F &gt; S</td>
<td>P &lt; 0.005</td>
</tr>
<tr>
<td>Vertex (11)</td>
<td>100–125</td>
<td>S &gt; F</td>
<td>n.s.</td>
</tr>
<tr>
<td>Vertex (11)</td>
<td>800–1075</td>
<td>F &gt; S</td>
<td>n.s.</td>
</tr>
<tr>
<td>Vertex (11)</td>
<td>1175–1400</td>
<td>F &gt; S</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

*The numbers in parentheses refer to the ROIs indicated in Fig. 4. inf. = inferior; sup. = superior; n.s. = not significant.
Fig. 7 Mean ± SEM (black curves and shading) source waveforms in the left and right frontoparietal subregions in fluent speakers (left) and stutterers (right) when the MEG signals were averaged with respect to mouth movement onset. Each division in the horizontal axis corresponds to 200 ms.

hemisphere the active areas concentrated closer to the hand area.

Figure 7 depicts the mean ± standard error of the mean source waveforms, averaged with respect to mouth movement onset, in the left and right frontoparietal ROIs in fluent subjects and stutterers. The activation increased slowly prior to movement onset bilaterally in the fluent speakers but was left-hemisphere dominant in the stutterers. The activation reached maximum at 69 ± 15 ms and 45 ± 9 ms in the left and right hemispheres of the fluent speakers and at 43 ± 12 ms and 65 ± 12 ms, respectively, in the stutterers (n.s. by hemisphere and group). When the difference between the right- and left-hemisphere waveforms was calculated in each individual, the stutterers showed significant left-hemisphere dominance, as compared with fluent speakers, at –550 to –450 ms [F(1,14) = 5.5, P < 0.04] and –250 to +25 ms [F(1,14) = 4.8, P < 0.05] with respect to mouth movement onset.

Responses averaged both with respect to word onset and mouth movement onset thus demonstrated a reduced time-locked activation in the right frontoparietal ROI of stutterers.

Modulation of cortical rhythms

Figure 8 illustrates task-related modulation of 20-Hz activity in one fluent speaker (C7) during a 15 s interval. Rhythmic activity over the central sulcus was suppressed after word onset and remained at a low level throughout the utterance. The 20-Hz TSE curves, depicting the mean amplitude of the oscillations with respect to word onset, illustrate that the suppression lasted for about 2 s and was concentrated over the central sulcus, with left-hemisphere dominance in this subject. Both the 10- and 20-Hz oscillations were suppressed by word onset and speech production in most subjects.

Generators of cortical rhythmic activity concentrated in two distinct loci along the central sulcus bilaterally, in the hand area and about 2 cm inferior to it along the central sulcus, approximately in the mouth area. As usual, cortex around the parieto-occipital sulcus produced rhythmic activity as well, with the dominant component around 10 Hz (Salmelin
Fig. 9 TSE curves of sources in the bilateral mouth and hand areas and around the parieto-occipital sulcus (see the schematic drawings of brains) across fluent speakers (black curves) and stutterers (grey curves), from 1 s before to 5 s after word presentation. Source strengths are given in nanoamperemetres (nAm). The question mark onset is indicated with a solid vertical line at 0.8 s. Mouth movement (M) and speech (S) onsets are marked with dashed vertical lines.

model ANOVA with two subject groups). The parieto-occipital 20-Hz activity, tested separately, did not show group differences.

In the mouth area, suppression of 20-Hz activity began 100–300 ms after word onset, i.e. ~700 ms before mouth EMG onset (Table 2). Onset latencies in the left and right mouth areas correlated with each other (fluent speakers $r = 0.69$, $P < 0.03$; stutterers $r = 0.70$, $P = 0.051$). The onset was earlier in the left than the right hemisphere in the fluent speakers but earlier in the right than the left hemisphere in the stutterers, in both the hand and mouth areas [hemisphere-by-group interaction $F(1,15) = 4.8$, $P < 0.05$]. Furthermore, the suppression started significantly earlier [$F(1,15) = 14.0$, $P < 0.002$] in the mouth than hand area in both hemispheres and subject groups. The onset latencies in the hand areas did not correlate with those in the mouth areas. The 20-Hz suppression onset or peak latencies did not correlate with mouth movement onset in either group, nor with the severity of stuttering.

For amplitude comparison, the TSE curves were integrated between −1000 ms and −200 ms to estimate the base level, between 200 and 2200 ms after word presentation to quantify the suppression, and between 2200 and 3200 ms to cover the post-suppression rebound above the base level which was detected in the left mouth area of some stutterers (Fig. 9). The base levels, suppressions, and rebounds were compared across hemispheres (left and right) and areas (hand and mouth; 2 × 2 mixed-model ANOVA with two subject groups).

The base levels did not vary between groups. However,
Table 2 Timing of the 20-Hz suppression in fluent speakers and stutterers in the hand and mouth areas of the left and right hemispheres

<table>
<thead>
<tr>
<th></th>
<th>Fluent speakers</th>
<th>Stutterers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mouth area (ms)</td>
<td>Hand area (ms)</td>
</tr>
<tr>
<td><strong>Left hemisphere</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Onset</td>
<td>120 ± 120</td>
<td>820 ± 200</td>
</tr>
<tr>
<td>Maximum</td>
<td>1100 ± 60</td>
<td>1480 ± 140</td>
</tr>
<tr>
<td>End</td>
<td>2440 ± 130</td>
<td>2140 ± 200</td>
</tr>
<tr>
<td><strong>Right hemisphere</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Onset</td>
<td>320 ± 130</td>
<td>840 ± 170</td>
</tr>
<tr>
<td>Maximum</td>
<td>1120 ± 50</td>
<td>1590 ± 170</td>
</tr>
<tr>
<td>End</td>
<td>2140 ± 160</td>
<td>2370 ± 190</td>
</tr>
</tbody>
</table>

The interhemispheric balance of suppression was significantly different in stutterers and controls [hemisphere-by-group interaction $F(1,17) = 4.5, P < 0.05$]. Comparison of the difference between right- and left-hemisphere suppressions in the hand and mouth areas ($2 \times 2$ mixed-model ANOVA for area $\times$ group) indicated that the interaction was due to suppression in the mouth area, which was stronger in the right than left hemisphere in stutterers but stronger in the left than right hemisphere in fluent speakers [$F(1,17) = 7.0, P < 0.02$, planned contrast]. The hemispheric imbalance of suppression tended to co-vary with severity of stuttering (right minus left; $\rho = -0.59, P = 0.09$, Spearman rank correlation, corrected for ties). A comparison of the differences between right- and left-hemisphere post-suppression amplitudes indicated a stronger rebound of 20-Hz activity in the left than right hemisphere mouth area in stutterers but not in fluent speakers [$F(1,17) = 7.1, P < 0.02$, planned contrast]. The hand area rhythms did not show significant group differences.

The opposite interhemispheric balance of 20-Hz suppression in the mouth area in fluent speakers and stutterers is also evident in Fig. 10A which illustrates the reductions as percent of the base level. The 20-Hz suppression was significantly stronger in the mouth than in the hand area [$F(1,17) = 12.1, P < 0.003$]. In controls, the difference in attenuation was pronounced both in the left and right hemisphere ($P < 0.02$, paired two-tailed $t$-test). However, in the stutterers the hand and mouth area suppressions did not differ significantly.

The locations of the dipolar sources representing the generators of the cortical rhythms in the hand and mouth areas are plotted in Fig. 10B. The hand area rhythms originated within 4 mm of the hand sensorimotor cortex, as identified in the finger movement task, similarly in both hemispheres; the hand area coordinates did not differ between groups. The mouth area rhythms originated about 20 mm inferior and 10 mm anterior to the hand area, along the central sulcus, in agreement with lip representation determined in previous intracranial and MEG recordings (Penfield and Rasmussen, 1957; Woolsey et al., 1979; Hari et al., 1993; Nakamura et al., 1998). In the left hemisphere, the mouth area sources were located on average 6 mm more inferiorly in stutterers than controls [$F(1,17) = 4.8, P < 0.05$, planned contrast, $2 \times 2$ mixed-model ANOVA for hemisphere $\times$ group] and, accordingly, the distance between the hand and mouth 20-Hz generators was significantly longer in stutterers than controls [30 versus 22 mm; $F(1,17) = 5.5, P < 0.04$, planned contrast].

The 10-Hz modulation was highly variable across subjects. The suppression was most salient in the parieto-occipital cortex. Modulation depths in both the hand and mouth areas were small (0.8–1.5 nAm or 6–12% of base level) and did
not differ from each other. Like in the 20-Hz range, the 10-Hz suppression started earlier in the mouth than in the hand area \( F(1,14) = 6.9, P < 0.02 \).

In both groups, reading words aloud was thus associated with pronounced modulation of the 20-Hz activity, particularly in the bilateral mouth areas. The suppression started well before mouth movement, and was correlated across hemispheres. Hand area suppression coincided with the vocalization prompt. The attenuation of rhythmic activity was stronger and earlier in the left hemisphere of fluent speakers but in the right hemisphere of stutterers. In fluent speakers, the modulation concentrated in the mouth area whereas in stutterers both hand and mouth areas were markedly engaged in speech production.

Discussion

Neuroimaging studies of reading

Reading familiar real words aloud is supposed to proceed via visual feature analysis, pre-lexical letter detection, and word-level visual processing to activation of the semantic and phonological representations, and finally to activation of the phoneme and articulation system to produce speech (Coltheart et al., 1993). Interaction between the adjacent processing stages is believed to take place at all levels. Whether these processing stages and their apparently modular structure, derived largely from studies on subjects with acquired language impairments, have actual counterparts in the organization of language in a healthy human brain is not clear. There is a general consensus of the involvement of at least the left-hemisphere perisylvian areas in reading, including the classical inferior frontal Broca’s and posterior superior temporal Wernicke’s areas, both from lesion and functional neuroimaging studies but the suggested correspondence between cortical areas and different subprocesses vary considerably according to the paradigm used (see, for example Poeppel, 1996).

Fiez and Petersen recently reviewed nine PET and fMRI data sets of reading aloud single words and reported consistent activations in the lower part of the motor cortex bilaterally (Brodmann area (BA) 4), supplementary motor area (BA 6), the inferior frontal cortex (Broca’s area, BA 44/45/insula), the superior temporal cortex bilaterally (BA 21/22/24, including Wernicke’s area on the left), left inferior occipitotemporal border (BA 37/19), left lateral occipital cortex (BA 18/19) and cerebellum (Fiez and Petersen, 1998). The possible functions were suggested to be visual analysis specific to word-like stimuli in the left inferior occipitotemporal region, semantic analysis near the border of the superior and middle temporal gyri, acoustically based phonological analysis in the left posterior temporal regions, articulatory based phonological analysis in the left inferior frontal cortex and insula, motoric aspects of language function in the motor cortex and supplementary motor area, and monitoring of one’s own voice in the bilateral auditory cortices. These interpretations were supported by both haemodynamic and lesion data.

On the other hand, Price, in her review of recent PET and fMRI studies on word comprehension and production, again paralleled by lesion data, concluded that semantic processing involves the left inferior frontal cortex (BA 47, anterior to Broca’s area), the left temporal pole (BA 20/28/38), and the posterior temporo-parietal (BA 39, angular gyrus) cortex (Price, 1998). Phonological retrieval was associated with activation of the left basal occipitotemporal border (BA 37/19) and the left frontal operculum (BA 44/45/insula). The left supramarginal gyrus (BA 40) was proposed to have a specialized role in converting the orthographic form of the word to the corresponding phonological representation. The two interpretations thus seem to vary most in the functional neuroanatomy of semantic processing (superior middle versus inferior anterior temporal lobe) and in the role of the left basal occipitotemporal cortex (visual analysis of word-like stimuli versus phonological processing).

The combined timing and localization obtained from MEG and intracranial recordings may shed some light on the possible roles of the different cortical areas. Letter-string specific neuronal responses, maximum at about 200 ms after stimulus presentation, were identified in the fusiform gyrus during intracranial recordings (Nobre et al., 1994). In MEG studies of word recognition, similar responses peaking 150–200 ms after word onset have been identified in the left inferior occipitotemporal cortex in fluent readers (Salmelin et al., 1996; Kuriki et al., 1998; Tarkiainen et al., 1999) but not in developmentally dyslexic subjects (Salmelin et al., 1996; Helenius et al., 1999). This early response may reflect an interface process which detects letter strings and conveys them from visual to language domain (Tarkiainen et al., 1999). Therefore, the role of this region, apparently corresponding to BA37/19 above, might be interpreted differently depending on the paradigm and the image subtractions used in the PET and fMRI studies.

By showing semantically constrained sentences, word by word, and varying the congruency of the final word in the sentence context, one can extract a response the strength of which increases for increasingly inappropriate sentence-ending words (N400 paradigm; Kutas and Hillyard, 1984). MEG studies have indicated that, based on this approach, the middle superior temporal cortex is involved in word and sentence comprehension 200–600 ms after word onset, with clear left-hemisphere dominance (Simos et al., 1997; Helenius et al., 1998). fMRI studies of sentence reading have also emphasized the role of the superior temporal cortex in reading comprehension (Just et al., 1996; Bavelier et al., 1997).

According to intracranial recordings, a semantic N400 response is also generated in the medial temporal structures and the temporal pole (Halgren et al., 1994; McCarthy et al., 1995; Nobre and McCarthy, 1995). Thus, MEG data would agree with the interpretation of Fiez and Petersen (Fiez and Petersen, 1998), whereas intracranial recordings in surgical patients would also support that of Price (Price, 1998).
Although mouth muscle and tongue movements tend to cause strong electric disturbances, MEG has been successful in imaging cortical activations associated with speech production. Both preparatory motor activity around the mouth and tongue areas (Sasaki et al., 1995; Kuriki et al., 1999) and activation of supplementary motor area and Broca’s area in a picture naming task (Salmelin et al., 1994) have been reported. To our knowledge there are no previous MEG studies of reading words aloud.

**Neuroanatomy of reading in the present data set**

In this study, we employed a delayed response paradigm which was successful in postponing the speech-related artefact, thus allowing us to image the spatiotemporal dynamics of word perception and preparation for overt output in a slightly extended time window. On the other hand, this somewhat artificial procedure could enhance a short-term memory component and promote verbal rehearsal of the word during the 800 ms delay from word onset to the vocalization prompt. The delay is, however, so brief that the subjects did not report any need to actively memorize the word during the process.

The overall spatiotemporal activation patterns in fluent speakers and stutterers agreed reasonably well with those in previous PET, fMRI and MEG studies reviewed above: (i) occipital and parieto-occipital responses starting at 100–150 ms after word onset (and continuing until the vocalization prompt), presumably involved in visual analysis; (ii) left and right inferior occipitotemporal clusters at 150–200 ms, with the left-sided sources probably reflecting letter-string specific analysis; (iii) left inferior frontal cortex (200–600 ms) and some activation in the homologous right-hemisphere locus, apparently reflecting articulatory aspects of phonological processing; (iv) left middle superior temporal cortex at 200–600 ms, as a signature of semantic activation; (v) left and right posterior parietal cortices (200–800 ms), possibly associated with phonological aspects of linguistic processing (left) or attentional aspects of visual perception (right; Mesulam, 1981; Nobre et al., 1997; Corbetta, 1998); and (vi) left and right motor and premotor cortices and supplementary motor area from 200 ms onwards, involved in motor preparation for oral output and actual vocalization.

The behavioural measures, i.e. mouth movement or speech onset latencies, did not differ between the groups. However, the cortical responses differed significantly in the left inferior frontal cortex and in the motor and dorsal premotor cortex bilaterally. Within 400 ms after seeing the word, the activation proceeded from left inferior frontal to motor cortex in the fluent speakers. The sequence was reversed in the stutterers who showed an abnormally early left motor/premotor response, followed by a delayed left inferior frontal activation. The exceptionally early motor cortical activation in stutterers may be reflected in the tendency to earlier initiation of mouth muscle activity in stutterers than in fluent speakers. During vocalization, 20 Hz activity was bilaterally suppressed in the mouth areas, with slight left-hemisphere dominance in fluent speakers and right-hemisphere dominance in stutterers. In stutterers, 20-Hz activity was also strongly suppressed in the hand areas. Moreover, stutterers failed to show a pronounced time-locked response in the right frontal cortex which was evident in the fluent speakers throughout vocalization. Below, we will discuss these differences in detail and relate the present findings with functional imaging and lesion data on fluent and dysfluent reading.

**Differences in time-locked activity after seeing the word**

Activation of the left inferior frontal cortex, ventral portion of BA44/45 and extending to the insula, has been reported in PET and fMRI studies of vocalized word reading (see Fiez and Petersen, 1998), verbal working memory (Paulesu et al., 1993; Fiez et al., 1996), memory for pitch (Zatorre et al., 1994), auditory and phonological processing (Fiez et al., 1995; Fiez, 1997; Gandour et al., 1998), and verbal fluency (Paulesu et al., 1997), and interpreted to reflect high-level articulatory encoding and involvement of a subvocal rehearsal system. Pugh and colleagues and Fiez and Petersen have proposed that the left frontal operculum contributes to the process of orthographic-to-phonological transformation from word to sound (Pugh et al., 1996; Fiez and Petersen, 1998). The timing in the present study is in line with these interpretations, as the articulatory activation in the left inferior frontal cortex started approximately at the same time as the semantic activation in the left superior temporal cortex, and both were preceded by the left posterior letter-string specific response. Naturally, our delayed reading paradigm could have evoked the subvocal rehearsal system as well. The processes listed above are all mutually intertwined and, for the present, we associate the left inferior frontal activation with articulatory encoding.

A severe disorder of articulation, without language disabilities, has been reported from focal lesions in Broca’s area (Mohr et al., 1978; Schiff et al., 1983) and in the lower part of the left precentral gyrus (Tonkonogy and Goodglass, 1981; Schiff et al., 1983; Mori et al., 1989). More recently, Dronkers identified a discrete region in the left precentral gyrus of the insula, damage to which was consistently associated with articulatory planning deficits (Dronkers, 1996). Broca’s aphasics may also show defective laryngeal control (Blumstein, 1995), a disturbance implicated in developmental stutterers as well (Freeman and Ushijima, 1978). Moreover, acquired stuttering, although in many ways behaviourally different from developmental stuttering (Koller, 1983), has been observed after lesions to Broca’s area and to the lower third of the premotor cortex (Tonkonogy and Goodglass, 1981; Freedman et al., 1984).

The cortical dynamics in our fluent speakers thus indicates
an early articulatory encoding process before activation of the bilateral dorsal motor and premotor cortices, presumably involved in preparation and performance of motor tasks (Freund, 1984; Wise, 1985). This sequence would seem quite reasonable for fluent speech. The reversed pattern in stutterers, on the other hand, suggests that motor programmes were initiated before preparation of the articulatory code, which could certainly result in disrupted speech and even stuttering. It is worth noting that defective articulation in early childhood is much more common in stutterers than in non-stutterers (cf. Bloodstein, 1995).

These early differences are quite short-lasting and could well be missed in haemodynamic recordings. However, the stutterers of both Wu et al. and Fox et al. displayed reduced blood flow in the inferior frontal cortex (BA 44/45/47) in the dysfluent versus fluent reading condition (Wu et al., 1995; Fox et al., 1996). Furthermore, Braun and colleagues found a more focal activation in stutterers than fluent speakers in this same region, both in fluency and dysfluency evoking tasks (Braun et al., 1997). The PET and MEG measures may reflect the same cortical disturbance in developmental stutterers.

**Differences in time-locked activity during speech production**

Based on lesion data, Ross and Mesulam proposed that the right perisylvian region plays a dominant role in affective prosody, i.e., in furnishing propositional speech with the melody and intonation appropriate to its emotional content (Ross and Mesulam, 1979). They further introduced the idea that the functional neuroanatomy of affective language in the right hemisphere mirrors that of propositional language in the left hemisphere, e.g., with motor aprosodia resulting from right frontoparietal lesions and sensory aprosodia from right posterior perisylvian lesions (Ross, 1981). According to Weintraub and colleagues, right-hemisphere damage may also cause deterioration of non-affective aspects of speech prosody (Weintraub et al., 1981). On the other hand, grammatically controlled (non-affective) prosody may be part of the articulatory encoding process in the left inferior frontal cortex (Sproat, 1995). The lesion data are supported by a recent fMRI study showing left-lateralized motor cortical activation during aprotodic speech but predominantly right-hemisphere activation during singing (Wildgruber et al., 1996).

Building the correct prosody is an essential part of fluent speech. Stuttering occurs most often in the beginning of a sentence or, more specifically, at borders between complete ideas to be expressed where the prosodic contour is determined (Starkweather, 1987; Karniol, 1995). Furthermore, stuttering occurs more frequently in an emotionally loaded than neutral context, and isolated words are rarely stuttered. Reading in chorus with another person, whispering, or following a pacing signal all reduce the need for self-initiated prosody, and they are all efficient in relieving stuttering symptoms (cf. Starkweather, 1987; Bloodstein, 1995).

In this framework, it is conceivable that the lack of right-hemisphere activation in stutterers, time-locked to stimulus or speech onset, is associated with difficulties in initiating the correct prosody. For the most severe stutterers, even the intonation of a single word could be problematic. The integration of words and prosody succeeds most of the time but its occasional disruption, in any part of the language network (Mesulam, 1990), could lead to dysfluency.

Fox et al. and Braun et al. found that the cerebral blood flow in the motor cortex was left lateralized in fluent speakers (Fox et al., 1996; Braun et al., 1997). In stutterers, the frontal activations were equal or right lateralized due to pronounced hyperactivity of the right motor cortex. The PET results were thus diametrically opposite to the time-locked MEG responses. However, the two measures may not be directly comparable: the MEG signal reflects synchronized firing in a large number of neurons whereas increase of blood flow may be associated also with diffuse non-synchronous neuronal activation (Hari and Salmelin, 1997).

**Modulation of 20-Hz activity in the hand and mouth areas**

The 20-Hz suppression in the mouth area, probably reflecting increased computational demands on the local neuronal populations in the bilateral mouth motor cortices (Salmelin et al., 1995; Hari and Salmelin, 1997), started well before mouth movement onset. Suppression of hand area 20-Hz activity has been reported during motor imagery of finger movements (Schnitzler et al., 1997). Similarly, the early 20-Hz attenuation may reflect mental preparation for speech production. During vocalization, the 20-Hz activity was suppressed in the hand area as well, indicating an extensive involvement of the precentral motor cortex beyond the mouth area.

In fluent speakers, the modulation was significantly stronger in the mouth than hand area, in line with previous reports on somatotopic organization of 20-Hz oscillations (Salmelin et al., 1995; Salenius et al., 1997). In the stutterers, however, 20-Hz activity in the hand area was relatively more suppressed, suggesting weaker segregation of the mouth and hand representations in stutterers than in controls. In principle, stutterers could have moved their hands as they prepared to speak, thus suppressing the rhythm. Although we did not observe overt hand movements, we cannot entirely exclude this possibility as hand EMG was not recorded.

The mouth motor cortex of fluent speakers showed both salient time-locked responses and marked attenuation of 20-Hz cortical activity. However, in the stutterers the right hemisphere displayed a pronounced 20-Hz suppression but essentially no time-locked responses. This dissociation suggests that the right rolandic region of the stutterers was
strongly active but the computations in the neuronal patches were not properly synchronized. Such lack of synchronization could arise from imprecise functional connectivity within this cortical area, which again may be associated with an incomplete segregation of mouth and hand representations in the motor cortex. Studies on rats have shown that in the development of vibrissae-related cortical fields the thalamocortical topographical maps are formed first and cortical connections between representations are established thereafter, within the first few postnatal days (Rhoades et al., 1996, 1997). If a similar process takes place in the human brain, a disturbance in the early postnatal organization of the neocortex could play a role in developmental stuttering.

The task-related suppression of rhythmic activity may correlate with haemodynamic measurements better than the short-lasting, time-locked responses. Combined EEG and PET recordings have indicated that when the posterior 10-Hz activity is attenuated the blood flow increases in the parieto-occipital region (Sadato et al., 1998). Assuming that a similar relationship holds for the 20-Hz activity, the stronger right- than left-hemisphere suppression in stutterers could be associated with higher blood flow in the right than left rolandic area. This is indeed what was reported in the PET studies of Fox et al. and Braun et al., even when the stutterers were fluent (Fox et al., 1996; Braun et al., 1997).

**Cortical correlates of stuttering**

Acquired stuttering has been reported following unilateral damage to the left inferior frontal cortex (Tonkonogy and Goodglass, 1981; Freedman et al., 1984) and to the left and right perisylvian and rolandic regions (Helm et al., 1978; Fleet and Heilman, 1985). Helm and colleagues associated transient acquired stuttering with left unilateral lesions and persistent acquired stuttering with bilateral damage (Helm et al., 1978). Fluent speech requires on-line merging of motor production, prosody, meaning and syntax (e.g. Benson, 1986). The left hemisphere seems to have a dominant role in semantic and syntactic processing and in the associated grammatical prosody, and the right hemisphere in control of affective prosody. Both hemispheres control the actual mouth, tongue and larynx movements. Integration of the bilateral language functions (Ross et al., 1997) is probably essential for fluent speech and could be disrupted in stuttering.

The present data suggest, in line with earlier neuroimaging data (Wu et al., 1995; Fox et al., 1996; Braun et al., 1997), that both hemispheres are affected in developmental stutterers whose dysfluency has persisted into adulthood. An entire language network, probably including articulatory and motor preparation for speech production and generation of correct prosody, is partly dysfunctional. The MEG data on fluent single word reading in stutterers thus seem to be in reasonable agreement with the theoretical framework proposed by Karniol, which emphasizes the complex interaction between articulation and prosody in stuttering (Karniol, 1995). Evidently, further detailed studies are needed for establishing the exact functional roles of the affected cortical areas, the influence of overt stuttering on the activation patterns, and the possible causal links between the distinct functional units.

**Acknowledgements**

We thank Lutz Jäncke for contributions in the initial phases of the project, Kai Lutz and Sven Langefeld for kindly performing the psychological testing of the subjects, and Riitta Hari, Päivi Helenius, Katri Kiviniemi, Matti Laine and Antti Tarkkainen for comments on the manuscript. This work was supported by the Academy of Finland (grant 32731), the Human Frontier Science Program and Volkswagen-Stiftung.

**References**


Single word reading in stutterers


