

HUMAN COMMUNICATION INVESTIGATED WITH MAGNETOENCEPHALOGRAPHY: SPEECH, MUSIC, AND GESTURES

Thomas R. Knösche, Burkhard Maess, Akinori Nakamura,
and Angela D. Friederici

Max-Planck-Institute for Human Cognitive and Brain Sciences
04103 Leipzig, Germany

- I. Introduction
- II. Processing Spoken Language
 - A. Recognition of Words and Voices—Early Preattentive Processes
 - B. The Processing of Syntax
 - C. The Process of Lexical Selection
- III. Processing Music
 - A. The Processing of Musical Syntax
 - B. Phrasing in Music
 - C. Sensory Motor Coupling
- IV. Perception of Gestures
- V. Conclusions
- References

In this chapter, we report a number of studies applying magnetoencephalography (MEG) to the investigation of the modes of human communication. We seek to elucidate the spatio-temporal organization of the processing of different levels of information during the perception of speech, music, and gestures. In particular, we investigated the preattentive processing of speech sounds, the processing of syntactic information in speech and music, the mechanisms of lexical selection during language production, the perception of musical phrase structure, the coupling between sensory and motor representations in musicians, and the understanding of emblematic hand signs. We applied source localization methods, enabling us not only to localize the active brain areas underlying the respective cognitive processes, but also to attribute precise timing to these processes. The actual choice of methods for both the physical modeling of the head (forward problem) and the reconstruction of the brain activity (inverse problem) has to be made anew in each case, carefully considering the nature and quality of the data, available supporting information (e.g., magnetic resonance images of the head), and the hypotheses on the possible outcome. Consequently, we applied a whole range of techniques throughout the studies reported in this chapter. The solutions to the forward model included simple spherical as well as

realistically shaped head models. Inverse methods used in our studies encompassed the reconstruction of focal sources (spatio-temporal dipole fit), of two-dimensional current distributions (brain surface current density mapping), and of three-dimensional voxel-based activity distributions (magnetic field tomography, multiple signal classification). Principal component analysis was used to disentangle different neuronal systems or to detect subtle effects obscured by irrelevant brain activity. Additional information from functional magnetic resonance imaging was also employed. The reported results highlight the ability of MEG to map the brain activity underlying high-level cognitive processes in space and time.

I. Introduction

The activity of the human brain involves a complex interplay of chemical and electrical processes. This complexity extends over both space and time dimensions. The principal means of transport, distribution, and integration of information are electrical impulses exchanged between the neurons and electrical potentials across the membranes of these neurons. This information processing is mediated by the activity of chemical substances (neurotransmitters, receptors, and so on) and accompanied by supporting physiological (e.g., metabolic) processes. The vast diversity of phenomena carrying information on brain functions naturally yields a great variety of possible means to extract this information. In animal research, very detailed investigations are possible (e.g., by electrical single-cell recordings or by mapping of neurotransmitters). In humans, the opportunities to watch the brain in action are more limited. In some cases, by-products of clinical diagnosis and treatment (e.g., intracranial electrical recordings in epilepsy patients) can be used. In general, however, less invasive methods have to be employed. These methods are characterized by the brain process they monitor, by their degree of invasiveness, and by their spatial and temporal resolution. Positron emission tomography (PET) and single-photon emission computerized tomography (SPECT) measure distribution and transport of radioactively marked chemical substances. They feature a reasonable spatial resolution (about 1 cm) and great variability of observable processes, due to a large choice of marker substances. However, the use of radioactive substances renders these technologies quite invasive and limits their use on healthy subjects. Functional magnetic resonance tomography (fMRI) makes use of strong magnetic fields, without any known side effects thus far. Its time resolution is considerably better than PET and SPECT (in the order of seconds instead of minutes). Moreover, it offers excellent spatial resolution (1 mm and less) and, therefore, has become very popular in brain research. All methods mentioned so far rely on metabolic and hemodynamic changes in the brain tissue, which indirectly reflect neuronal

activity. A much more direct measure of the activity of nerve cells is delivered by electroencephalography (EEG). The potential differences on the head surface are mostly caused by postsynaptic potentials of neurons (Niedermeyer and Lopes da Silva, 1999). However, they are modified considerably by the electrical properties of the head tissues. EEG yields an excellent time resolution (fractions of a millisecond), but only very poor spatial specificity.

Magnetoencephalography (MEG) monitors the same electrical brain activity as EEG. The magnetic field outside the head is determined by current flows caused by postsynaptic potentials within the neural tissue. Whether the principal spatial resolution of MEG is superior to the one of EEG remains a matter of debate. A decisive answer to this question is difficult, if not impossible, since resolution depends on a number of factors, including adequate spatial sampling (number and arrangement of measurement channels) and signal-to-noise ratio. It also greatly depends on the use of model-based algorithms to disentangle the influences of the various active brain areas onto the signals, the so-called source localization techniques.

Why should one use MEG, which is quite an expensive technology, at all? First, although the principal spatial resolution might be the same, the influence of the extracranial head tissue makes accurate modeling of the generators of EEG quite difficult, while for MEG, practically only the interior of the skull has to be considered (Hämäläinen and Sarvas, 1987). Modeling is further facilitated by the fact that in MEG the exact positions of the sensors with respect to the head are much easier to determine. Second, MEG is not yielding the same information as EEG. That is, if MEG is measured additionally to EEG (simultaneously or in separate sessions), extra information on the intracranial processes can be won, which is not obtainable by simply adding extra EEG electrodes.

Brain imaging methods can deliver answers to the question: which parts of the brain are active at which time? It is important to note that even a perfect answer to this question does not guarantee the understanding of how the brain is functioning. Only clever experimental design and sound theoretical reasoning, together with reliable brain imaging, can shed light on the complex processes that make us think. Naturally, brain imaging was first applied to processes, which are clearly circumscribed and easily elicited by external stimulation or instruction. These processes include the primary sensory and motor activations. However, in recent years it became evident that also the most complex mental phenomena (e.g., language, decision-making, and even emotions) can be mapped in space and time.

There are a great variety of methods to extract this spatio-temporal information from MEG (and EEG) measurements. The so-called neuroelectromagnetic inverse problem can be split into two relatively independent sub-problems: (1) the way to account for the physical properties of the surrounding head tissue (forward problem), and (2) the modeling of the source using reasonable assumptions and

prior knowledge, leaving only as many free parameters as can be determined by the measured MEG/EEG data (inverse problem).

As solutions for the *forward problem*, available techniques range from very simple, such as spherical head models, to quite complex, such as inhomogeneous and anisotropic head models based on the finite element method (for an overview, see, e.g., Zanow, 1997). The actual choice depends on the accuracy requirements of the problem, the availability of anatomical information (e.g., individual MRI), certain practicability issues (e.g., the finite element method [FEM] is computationally very costly and cannot always be used), and modeling peculiarities (e.g., patients with holes in the skull or infants with open sutures). In some of our earlier work (Friederici *et al.*, 2000a; Groß *et al.*, 1998; Knösche *et al.*, 1999), we used a spherical head model or the boundary element method (BEM) (Hämäläinen and Sarvas, 1987) with a standardized head shape. In later studies, we introduced greater accuracy by individually scaling the standard head model to subject's head shape information, without the need to employ individual MRI information (Haueisen and Knösche, 2001; Knösche *et al.*, 2002, in press; Maess *et al.*, 2001, 2002). The very complicated brain networks revealed by our study on hand sign recognition (Nakamura *et al.*, 2004) required the use of BEM models based on anatomical information from individual MRIs. In the near future, even more elaborate head models based on FEM (see, e.g., Miller and Hendriquez, 1990) will be used for greater accuracy and special cases (holes in the skull, and so on). The principal problems for the use of FEM in source localization, namely the great computational costs and the robust handling of singularities, are currently tackled and very promising initial results were obtained (Wolters *et al.*, 2002, in press).

The techniques to solve the *inverse problem* are numerous and diverse (the following description is focused on techniques used in our MEG studies and, therefore, is not complete; for an overview, refer to, e.g., Knösche, 1997). The choice of a particular method depends on the nature of the expected results (e.g., few focal sources or extended widespread activity), the availability of additional information (e.g., locations of activated areas from fMRI, shape of individual cortical sheet), the signal-to-noise ratio of the data, and so on. The spatio-temporal multiple dipole model (Scherg and Berg, 1991) is one of the most popular approaches. It is based on strong assumptions: the generators of the MEG or EEG are described by a small number of focal centers of activity, which remain active for some time. If these assumptions are physiologically adequate, the model offers excellent data reduction (e.g., the time courses of for example 148 MEG channels are reduced to the time courses, positions, and orientations of two or three current dipoles). This model was successfully employed by our group to localize the centers of preattentive auditory discrimination (Knösche *et al.*, 2002), as well as initial parsing of syntactic structures in speech (Friederici *et al.*, 2000a) and in music (Maess *et al.*, 2001).

However, there are a number of possible reasons, why the current dipole model might not be chosen. First, the focality assumption can be violated to an intolerable extent. Second, the number of sources could not be known in advance with sufficient certainty. Third, the signal-to-noise ratio could be too low. In such cases, methods that reconstruct a distributed current source density distribution on the brain surface (e.g., brain surface current density [BSCD]) (Knösche *et al.*, 1996) can be applied. Such a method delivers a blurred picture of the brain activity, which is projected onto the brain surface. The depth information is sacrificed on purpose, because it is especially ill-determined by the data alone, if no plausible assumption on the shape of the sources (e.g., focality) is possible. However, based on probabilistic arguments, three-dimensional reconstruction techniques are also proposed. One such method is magnetic field tomography (MFT) (Ioannides *et al.*, 1990). We used the BSCD method to reconstruct the generators of preattentive auditory processing (Lattner *et al.*, 2003) as well as sensory-motor coupling in pianists when they listen to piano music (Haueisen and Knösche, 2001), and the MFT method for localizing brain regions, which are responsible for early syntactic processing (Gross *et al.*, 1998). The BSCD method, as any other distributed source reconstruction method, can yield a quite complex and confusing picture of activity, which results from the overlap of many different neuronal networks, each of which may or may not reflect different functions in a complex cognitive process. In order to disentangle functionally different components of the reconstructed activity, factor analysis (e.g., based on principal components analysis [PCA]) is a suitable tool. Each of the resulting components is characterized by a topographic map (on the brain surface) and a time course for each of the experimental conditions. The topographies are easier to attribute to certain brain systems than the original BSCD map and the time courses can be tested statistically for influences of the experimental variables. This technique allows the extraction of very subtle effects from overwhelming common activity (Maess *et al.*, 2002) or the disentanglement of very extensive networks of brain areas involved in a certain task (Nakamura *et al.*, 2004).

Another class of techniques are the so-called *scanning methods*. They are not true inverse algorithms, as they do not reconstruct the brain activity underlying a particular measurement as a whole (Vrba and Robinson, 2001). Instead, they consider each point or region in the brain separately and assess the plausibility of this point/region being a contributor to the measured activity. Instead of one possible generator distribution, which can explain the measured data, one obtains a collection of possible contributors (i.e., the description of an entire class of solutions). This allows the extraction of sensible information on the localization of the generators, even if true inverse methods fail. Scanning methods include: the *synthetic aperture magnetometry* (SAM) (Vrba and Robinson, 2001) and the *multiple signal classification* (MUSIC) (Mosher *et al.*, 1992). The MUSIC method was successfully applied by our group for the localization of the generators of the

closure positive shift (CPS) in music, which marks the processing of musical phrase boundaries (Knösche *et al.*, 2005).

In this chapter, we focus on the investigation of the modes of human communication. Questions on the spatio-temporal organization of the processing of different levels of information during the perception of speech, music, and emblematic hand signs will be asked and answered using MEG, partly in combination with other methods, like EEG and fMRI.

II. Processing Spoken Language

The most fundamental means of information exchange between humans is speech. The fast and accurate processing of speech signals is crucial for our ability to live together in society. The cortical representation of these capabilities have been investigated for more than a century, starting with Paul Broca's (1861) and Carl Wernicke's (1874) famous work. In recent years, modern neurophysiological and brain imaging techniques have provided access to more detailed information on the different processes subserving the perception and understanding of speech. Friederici (1995, 1998, 2002) proposed a model of the different stages of speech processing (Fig. 1).

After initial acoustic processing in Heschl's gyrus and the anterior part of the planum temporale (Lütkenhöner and Steinsträter, 1998), as reflected by the event-related potential (ERP) components *P50* and *N100* (*P50m* and *N100m* in MEG), higher acoustic analysis occurs in the planum temporale (BA42), including phoneme identification and voice identification (Griffith and Warren, 2002; Jäncke *et al.*, 2002; Knösche *et al.*, 2002). PET studies by Démonet *et al.* (1992) and Zatorre *et al.* (1996) have shown that phonetic segmentation and sequencing involves the posterior-superior part of BA44 (not the inferior part known as Broca's area). The superior temporal gyrus seems to support the identification of the final word form and the word category. This information is necessary for the construction of syntactic structures, which appears to involve Broca's area and its right hemisphere homologue as well as bilateral planum polare, as has been suggested by source localization based on the related electrophysiological component, the *early left anterior negativity* (ELAN/ELANm) (Friederici *et al.*, 2000a; Gross *et al.*, 1998; Knösche *et al.*, 1999). Later processes are mainly reflected by specific ERP components and their MEG counterparts, like the *N400* reflecting lexical-semantic integration (Kutas and Hillyard, 1980), the *P600* signaling syntactic re-analysis, and the *left anterior negativity* (LAN), which is interpreted as an index of thematic role assignment (Rösler *et al.*, 1993). The following sub-chapters deal with specific aspects of this processing chain. First we describe work on early phonological processes, in particular the recognition of word and voice

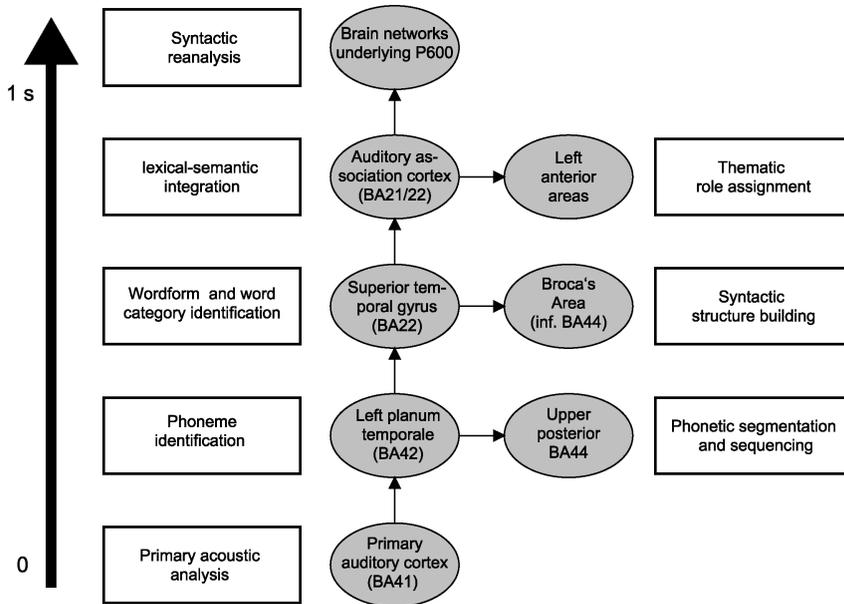


FIG. 1. Neurocognitive model of speech comprehension.

information. Second, we present a series of studies attempting to localize brain mechanisms supporting early syntactic structure building. Finally, aspects of semantic categorization are investigated in an overt picture-naming task.

A. RECOGNITION OF WORDS AND VOICES—EARLY PREATTENTIVE PROCESSES

The auditory cortices (primary auditory cortex and auditory association cortex) are known to harbor a surprisingly extensive arsenal of capabilities, described as “primitive intelligence of the auditory cortex” (Näätänen *et al.*, 2001). These processes generally occur approximately within the first 200 ms after stimulus onset and discriminate both basic physical (e.g., pitch, duration, volume) and higher phonetic (e.g., recognition of phonemes and voices) properties of the stimulus. In electrophysiological experiments (EEG, MEG), such processes are reflected by a number of typical waves (e.g., P50, N100, P200). Of special importance is the so-called *mismatch negativity* (MMN/MMNm), which is elicited between 100 and 200 ms after stimulus onset by any discriminable change of a repetitive aspect of the auditory input (Näätänen, 2001; Näätänen *et al.*, 1978). This effect has been extensively investigated and proven to reflect preattentive information processing, including the processing of phonemic

information (Näätänen *et al.*, 1993; Sams *et al.*, 1985). Its origin has been localized in the auditory cortex (Giard *et al.*, 1995; Hari *et al.*, 1984).

One important, socially relevant, type of information in speech input is the speaker's voice. Voice conveys information on the speaker's identity, gender, age, mood, and so on. It is rather independent of phonemic information. The early preattentive processing of voice information has been demonstrated by Titova and Näätänen (2001) using the MMN paradigm and is supported by fMRI data showing voice-specific brain areas in the auditory cortex (Belin *et al.*, 2002). However, from these studies it is not clear, whether only unspecific differences between the physical properties of the stimulus and the respective short-term memory traces are reflected, or if long-term memory traces of voice prototypes play a role. In order to clarify this issue, we designed an experiment involving prototypical (human) and non-prototypical (computer manipulated) voices. A prototypical male voice was used as a frequent standard stimulus, while three different versions of a female voice were presented as rare deviants: one prototypical, one computer-manipulated (pitch was shifted to match the male voice), and one completely computer-synthesized voice, matching the female voice in its main parameters. This way, the pitch-shifted female voice was physically more similar to the standard (differing in formant structure only) than the natural female voice (differing in both pitch and formant structure). With this stimulus material, an MEG experiment was performed (Lattner *et al.*, 2003) and the observed mismatch responses were projected onto the brain surface using a minimum norm least squares algorithm (Hämäläinen and Ilmoniemi, 1994; Knösche *et al.*, 1996, 1999; Wang *et al.*, 1992). If the mismatch response depended on short-term memory traces only, it should reflect the physical difference of the deviant and the standard, hence the pitch-shifted female voice should evoke the smallest response, the natural female voice should induce a larger response, and the computer-generated female voice should evoke the largest response. If, on the other hand, long-term memory traces of voice prototypes would play a role, the responses might mirror the stimulus prototypicality instead (i.e., the natural female voice shows the smallest MMNm, followed by the manipulated and the computer-generated voices). The results (Fig. 2) clearly demonstrate that the amplitude of the mismatch response is mainly modulated by the stimulus prototypicality. Hence, the prototypicality of speech sounds (i.e., their agreement with typical patterns of human voices as stored in long-term memory) affects the early preattentive processing stages within 200 ms after stimulus onset. The remarkable independence of the mismatch response of the physical stimulus properties (not the pitch-matched but the naturally sounding voice produced the smallest response) leads to the conclusion that the different acoustic features are not separately compared between standard and deviant. Instead, it seems that the mismatch response reflects a "gestalt-like" processing of voice information.

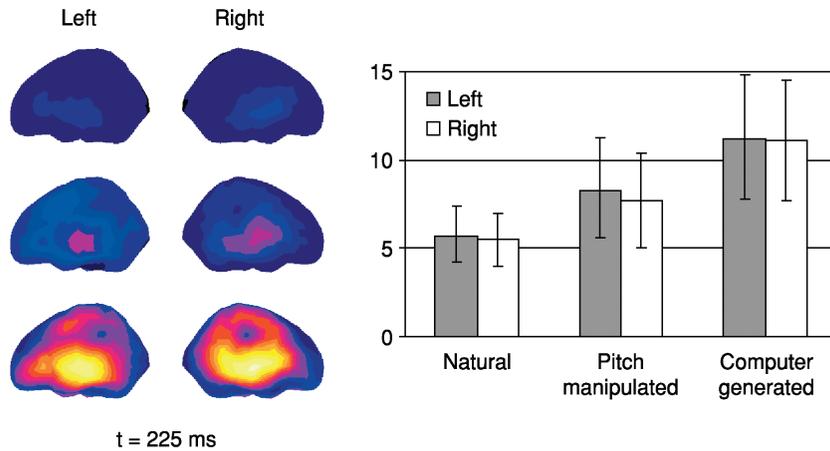


FIG. 2. Brain surface current density maps in response to the natural female voice (top row), the pitch manipulated female voice (middle row), and the computer generated voice (bottom row). The color scale runs from black ($0 \mu\text{Am}/\text{m}^2$) over blue, red, and yellow, to white ($4 \mu\text{Am}/\text{m}^2$). The bar chart indicates the average current densities per hemisphere in a time window between 100 and 300 ms after stimulus onset. The described differences between the various deviants were statistically significant, while there was no hemisphere effect (adapted from Lattner *et al.*, [2003], Fig. 3, with permissions of John Wiley and Sons Ltd).

Not only voice information, but a large number of aspects of auditory input, both basic acoustic and higher cognitive, have been shown to be processed preattentively and to be reflected by the mismatch response (for an overview, see Näätänen, 2001), including phonemic information (Näätänen *et al.*, 1997; Winkler *et al.*, 1999). In contrast to voices, phonemes are linguistic entities and, therefore, may be stored in long-term memory in a different way (symbolic). Phoneme and voice information (together with many other aspects) have to be extracted from the auditory input in a relatively short time. This raises a number of questions: (1) Are there different neural substrates for the preattentive extraction of different kinds of information, in particular between linguistic (phonemes) and extralinguistic (voices) information? (2) Are the different kinds of information processed serially or in parallel? For the parallel case, there could be an *independent parallel processing* (each feature processed and compared independently and eliciting its own MMN) or an *integral parallel processing* (different parallel processes “gather evidence” for deviancy, until a threshold is reached and detected by some central mismatch unit, which elicits the MMN). Different models generate a different prediction about the MMN. If serial processing is assumed, the amplitude and latency of the MMN due to a feature conjunction should be equal to one of the features presented in isolation. Independent parallel processing implies that the MMN responses of the separate features would simply add up. If

the features within a feature conjunction are processed in an integrative parallel way, the MMN amplitude would not be affected, but the speed and reliability of the deviancy detection process would be enhanced. This would lead to a shorter and less variable latency of the MMN. These hypotheses were put to the test in an MEG experiment using an oddball paradigm (Knösche *et al.*, 2002). Besides the frequent standard, three different kinds of rare stimuli were presented, deviating in phoneme information, in voice information, or in both at the same time. The first negative wave after stimulus onset, called N100m, was evaluated. It was larger in amplitude for the deviant compared to the standard conditions, indicating the expected mismatch response (MMNm). Dipole localization was carried out for this component. The equivalent dipoles were slightly but significantly more anterior for the deviant conditions (Fig. 3). This is consistent with the simultaneous existence of a N100 source and a more anterior MMN source (Csepe *et al.*, 1992). Between the various deviants no differences in amplitude or dipole location could be detected. This is somewhat surprising, as the exact location of the MMN source has been found to be a function of the dimension of change in a number of studies (Alho *et al.*, 1996; Diesch and Luce, 1997; FrodlBauch *et al.*, 1997; Giard *et al.*, 1995). However, in most cases these differences were small (<1 cm) and it can be shown that such differences could have been missed by the statistical power at hand.

Moreover, hemispherical differences in dipole location were found (more anterior and lateral in the right hemisphere) and can be explained by systematic

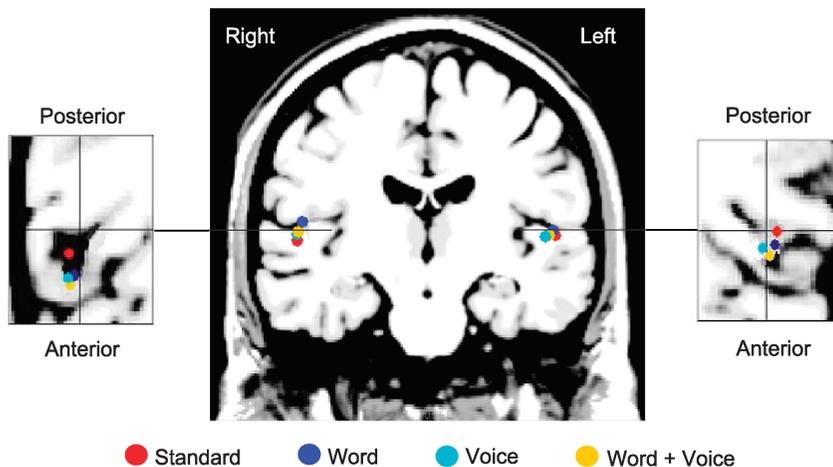


FIG. 3. Average locations of the generators of the N100m/MMNm, superimposed to coronal slice of standard MRI. Small insets show location in axial slice (adapted from Knösche *et al.*, [2002], Fig. 2, with permission of Elsevier).

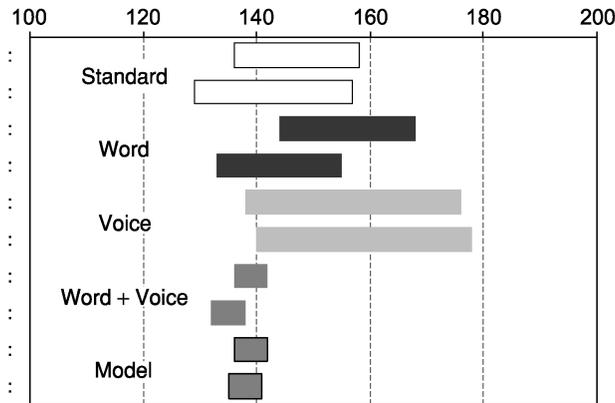


FIG. 4. 95% confidence intervals of N100m/MMNm latencies. The lowermost bar shows the mathematical prediction of an integral parallel model for the conjunction of phonetic and voice features (adapted from Knösche *et al.*, [2004], Fig. 4, with permission of Elsevier).

anatomical asymmetries (see, e.g., Penhune *et al.*, 1996). The amplitude of the response was stronger in the left hemisphere. This effect was equally present in the standard and all deviant conditions. Left lateralization of both N100 (Szymanski *et al.*, 2001) and MMN (Alho *et al.* 1998, Rinne *et al.*, 1999; Shtyrov *et al.*, 1998, 2000) for complex speech sounds has been reported in literature¹.

However, the most important question to be answered concerned the mechanism of the processing of simultaneous deviancy in the voice and phoneme domains. Since the amplitudes did not differ between the simple and combined deviant conditions, the independent parallel model can be excluded. Additionally, the latencies (Fig. 4) clearly demonstrate that the deviant detection based on a feature conjunction is faster (shorter latency) and more stable (less variance in latency). This observation is consistent with the integral parallel model. A precise mathematical model can be created, which assumes separate units (neural networks) for each stimulus dimension and extracts the respective feature to compare it to memory traces. A central unit integrates all of the differences and, if a certain threshold is reached, starts a signaling process, which is marked by the MMN or MMNm. The prediction of such a model is depicted in the lowermost bars in Fig. 4. These effects on the MMNm latency are in line with the observed redundancy gain effects in speeded classification experiments, where

¹Note that 9 out of 11 of the subjects in this study were males. There is some evidence that language related processes are more left lateralized in males than in females. See also subsection 2. The Processing of Syntax.

the classification of a fixed combination of words and voices was classified faster and with less jitter in reaction time, compared to the case, where both features varied independently (Mullenix and Pisoni, 1990; Wood, 1974). Hence, the MMNm seems to reflect the neural basis for those behavioral effects. It could be speculated that this increase in speed and reliability by the simultaneous change of more than one feature is a general property of the change detector mechanism underlying the MMN/MMNm.

B. THE PROCESSING OF SYNTAX

After phonemes have been identified and sequenced, words are categorized and recognized, presumably in the superior temporal gyrus (BA22). Once the word category information is present, the initial syntactic structure can be established. Electrophysiological evidence suggests that the integration of a new word into such a structure occurs between 150 and 300 ms after the onset of the respective word. If the word category does not match the expected one, normal initial structuring is not possible and a characteristic ERP component with a negative maximum at frontal electrodes is elicited, called *early left anterior negativity* (ELAN) (Friederici, 1995, 1998; Friederici *et al.*, 1993). Although the lateralization of this component can vary considerably between subjects, there seems to be a tendency for a left hemispherical dominance.

Although the distribution of the ELAN suggests generators in the anterior temporal or frontal cortex, a precise and reliable localization is difficult. Possible candidates for the underlying network are delivered by brain imaging studies, which identified superior temporal and inferior frontal regions (Caplan *et al.*, 1998; Just *et al.*, 1996; Mazoyer *et al.*, 1993; Stromswold *et al.*, 1996). However, these studies cannot decide, if the localized brain areas are involved in the early (initial structure building) or later (re-analysis) stages of syntactic processing, nor can they answer the question as to which of these areas accommodate the generators of the ELAN.

For this purpose, a series of MEG experiments was performed. The paradigm and material was similar to those used by Friederici *et al.* (1993) and Hahne (1998). German sentences were constructed, ending with a participle, which served as the critical item. This word either followed the auxiliary verb in a grammatically correct fashion, or it was preceded by a pronoun, causing a violation of the word category expectation. Appropriate filler sentences eliminated any predictability of the syntactic information.

In a first study, a 148-channel whole-head MEG system (MAGNES WHS 2500, 4D-Neuroimaging Inc., San Diego, CA) was employed (Knösche *et al.*, 1999). Differences in the MEG responses to the critical word between syntactically

incorrect and correct sentences were found for two distinct time windows: about 100 ms to 200 ms (ELANm) and about 300 ms to 500 ms after word onset. *Brain surface current density* maps were computed and integrated over regions of interest for statistical treatment (10 subjects). It could be shown that both frontal and temporal areas of both hemispheres are involved in the processing of the syntactic violation between 100 and 200 ms after onset of the critical word. There was no dominance of the left hemisphere. This result was largely confirmed by a second study using a BTi TWIN MAGNES system (2×37 channels) (Gross *et al.*, 1998). This time, *magnetic field tomography* (Ioannides *et al.*, 1990) was employed to obtain three-dimensional maps of statistical differences between correct and syntactically incorrect conditions for five subjects (see Fig. 5). Both inferior frontal and temporal areas in both hemispheres were found, with considerable interindividual differences. This time, there was a clear

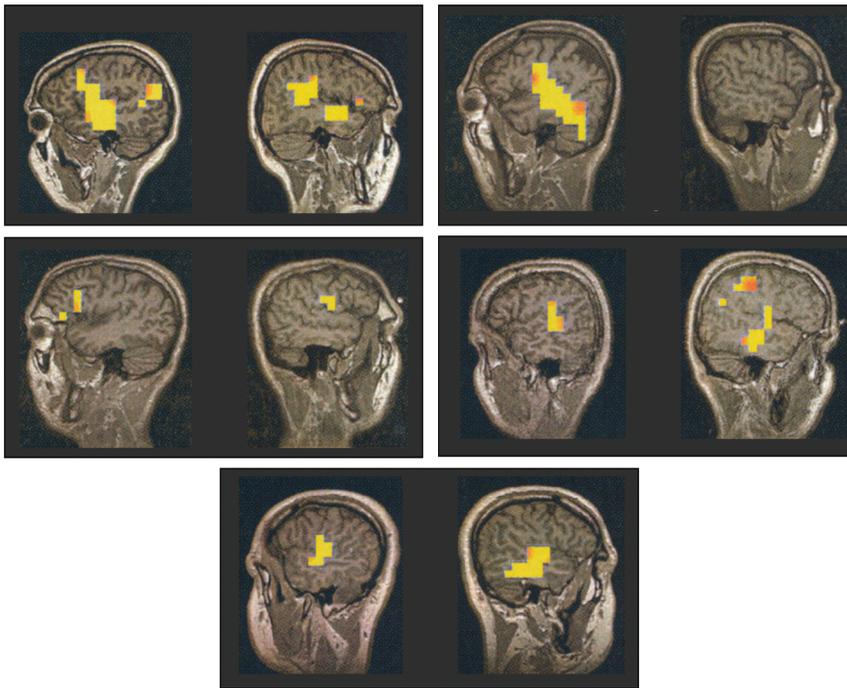


FIG. 5. T-test maps for magnetic field tomography results between syntactic correct and incorrect conditions for the latency of the ELAN. Colored voxels denote $p < 0.00001$ (uncorrected for multiple comparison) (adapted from Gross *et al.*, [1998], Fig. 4, with permission of Springer Science and Business Media).

dominance of the left hemisphere in both inferior frontal and superior temporal areas. In an attempt to obtain more reliable and precise results, a third study was conducted with a higher number of items presented to each subject (390 per category, compared to 130 in the other two studies; Friederici *et al.*, 2000a). Again, the syntactically incorrect sentences elicited an ELANm component between 100 and 200 ms after the onset of the violating word. The MEG data in a 20-ms time interval around the peak of the ELANm were subjected to a dipole fit procedure. Four dipoles were put at initial seed points, which were taken from an fMRI study with a similar paradigm (Meyer *et al.*, 2000). They were located in the inferior frontal (pars triangularis) and anterior temporal (planum polare) cortices of both hemispheres. During the fitting procedure, the dipoles were constrained to stay within 10 mm of the seed points. See Fig. 6 for the results. The obtained model explained on average 91% of the variance of the data. In four out of five subjects, left hemispherical activity was clearly stronger.

These MEG results clearly indicate that the ELAN is generated by a temporo-frontal network, involving inferior frontal (Broca) as well as anterior temporal (planum polare) areas in both hemispheres. This result is in line with results from the brain imaging and lesion literature. Inferior frontal cortex was shown to be involved in the processing of syntactically complex sentences by a number of studies (Caplan *et al.*, 1998; Friederici *et al.*, 2000b, 2003; Just *et al.*, 1996; Stromswold *et al.*, 1996). The direct involvement of this area into the generation of the ELAN was demonstrated in patients with left anterior cortical lesions (Friederici *et al.*, 1998, 1999). On the other hand, the contribution of temporal regions to the processing of syntax is also documented by brain imaging experiments (Friederici, 2002; Friederici *et al.*, 2000b, 2003; Mazoyer *et al.*, 1993; Meyer *et al.*, 2000) and lesion studies (Dronkers *et al.*, 1995; Friederici and Kotz, 2003).

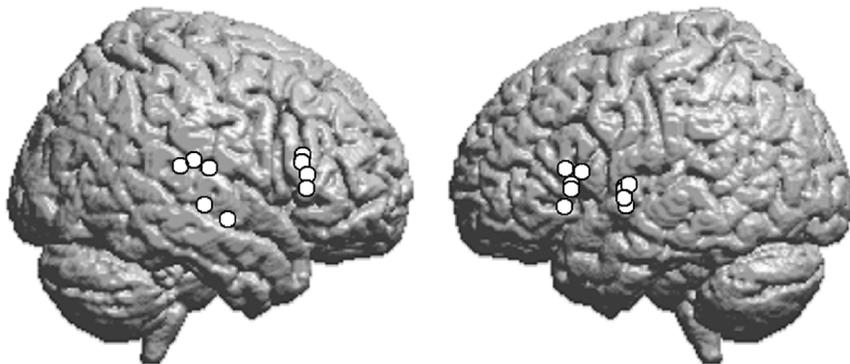


FIG. 6. Dipole locations (white circlets) for the syntactically incorrect condition at the latency of the ELAN, overlaid to a standard MRI-rendered brain surface.

The question of lateralization of the early syntax-related activity is not unanimously answered by the aforementioned studies. However, it is striking that Gross *et al.* (1998), who used male subjects only, found clear left hemisphere dominance, while Knösche *et al.* (1999), employing females only, found a more bilateral distribution. In the study of Friederici *et al.* (2000a), the three male subjects exhibited a left hemisphere dominance, while of the two female subjects, one was left and the other was right hemispheric. Such gender-specific lateralization of language processes has been found by some studies using lesions (McGlone, 1978), dichotic listening (Lake and Bryden, 1976), and brain imaging (Jaeger *et al.*, 1998).

Summarizing, there is evidence that early syntactic processing is supported by a brain network involving inferior frontal and anterior temporal regions in both hemispheres. This network houses the generators of the ELAN and ELANm. There is a tendency for stronger activity in the left hemisphere for most males, but usually not for females.

C. THE PROCESS OF LEXICAL SELECTION

For the production of spoken words, a number of mental operations are necessary. Levelt *et al.* (1999) list the stages of this process, which are assumed to occur in a serial fashion. The *object recognition* and the *conceptual preparation* phases are followed by the *lexical selection* process, *phonological code retrieval*, *phonological* and *phonetic encoding*, and finally the actual *articulation*. A second important component of this process, however, is a constant self-monitoring, which is based on phonological word forms and works back into earlier stages of the process, in particular on the conceptual phase (Wheeldon and Levelt, 1995). In the remaining part of this section, we deal with the *lexical selection* operation. Behavioral studies on picture-naming have estimated the time window for this process roughly around 150 to 275 ms after the stimulus onset (Levelt *et al.*, 1998). Although PET and fMRI experiments on speech production cannot directly distinguish the different stages of speech production due to their low temporal resolution, a meta-analysis of 58 such studies by Levelt and Indefrey (2000) could, by systematic comparison of the different paradigms, shed some light onto the brain regions involved in each of the different subprocesses of speech production. They identified the middle portion of the left middle temporal gyrus as the relevant area to support the lexical selection process.

In order to link these lines of evidence concerning timing and the neuronal substrate, an MEG experiment was designed (Maess *et al.*, 2002). As paradigm, the well-established *semantic category interference effect* was used (Kroll and Steward, 1994). When items (pictures) appear within lists of items of the same semantic category (e.g., a horse among other animals), reaction times are 30 to 40 ms

longer compared to items in lists of mixed categories. This effect was shown to be lexical rather than conceptual (Damian *et al.*, 2001). It is explained by the coactivation of semantically related concepts through spreading activity in the semantic network, and the subsequent competition of these activated concepts for lexical selection (Levelt *et al.*, 1999).

In our MEG experiment (Maess *et al.*, 2002), line drawings of well-known objects belonging to five categories (animals, vehicles, tools, clothing, furniture) were presented in blocks, which consisted either of objects of a single category or of objects of all categories in a mixed fashion. The subjects were instructed to overtly name the objects. Reaction times confirmed the presence of the *semantic category interference effect*, as subjects were 26 ms faster in the mixed-category condition.

During the experiment, MEG data were collected using a 148-channel whole-head magnetometer system (MAGNES WHS 2500, 4D-Neuroimaging Inc., San Diego, CA). Data were averaged and BSCD maps were computed (Knösche *et al.*, 1999). Factor analysis based on principal components was applied to the BSCD maps² and the time courses of the factors were compared statistically between conditions. Only one of the factors exhibited significant differences between the same- and mixed-category conditions. Its topography suggests left temporal centers of activity (Fig. 7). Significant differences between conditions were observed in an early (150 to 225 ms) and a late (450 to 475 ms) time window (Fig. 7). For the early time window, the result is in striking agreement with behavioral studies with respect to the timing (Levelt *et al.*, 1998), and with brain imaging work with respect to the localization in left temporal regions (Levelt and Indefrey, 2000). The second window of differential activity between same- and mixed-category conditions (450 to 475 ms) is attributed to the self-monitoring process, which might be located in left temporal cortical areas as well (Levelt and Indefrey, 2000).

III. Processing Music

Although generally not directly linked to the exchange of information, music plays an important role in human communication as it expresses cultural and social identity as well as emotion and, to a certain extent, semantic contents. It is a very old and uniquely human faculty.

²Source position formed one dimension of the decomposed matrix, while time instant (between 150 and 500 ms), condition, block and subjects were arranged along the other dimension. Following cutoff after nine factors, PROMAX rotation was applied.

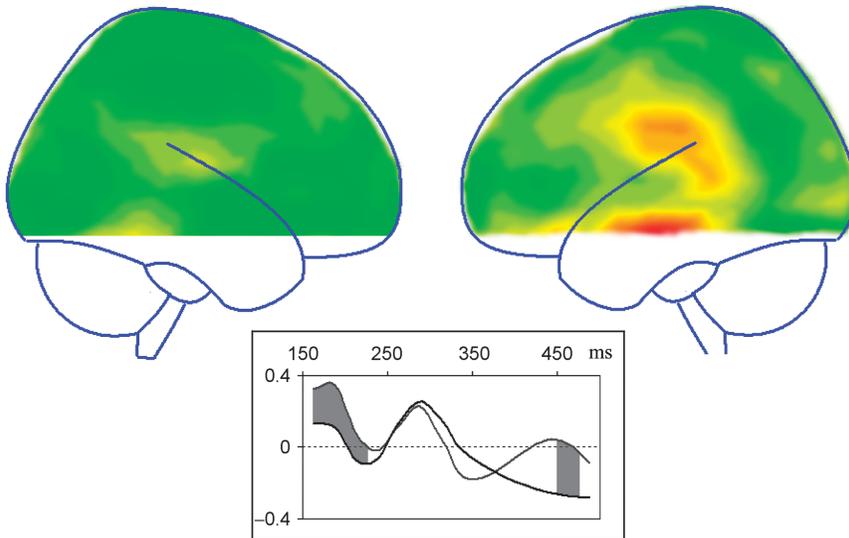


FIG. 7. Brain surface current density map of 6th PCA factor of the semantic category interference effect. Red areas indicate high and green ones low activity. Dark line = same-category condition, grey line = mixed-category condition (adapted from Maess *et al.*, [2002], Fig. 4 and 5, with permission of MIT Press).

Music shares many traits with language, but there are also fundamental differences. The degree of similarity between music and speech has often been a matter of debate. Undoubtedly, both consist of streams of sound elements, where principles of grouping and segmentation apply. The sound elements themselves are characterized by the same parameters: pitch, duration, articulation, timbre, and loudness, although their dynamic ranges and resolutions are certainly different between music and speech. Furthermore, the sequential structure of the sound elements in both domains is subject to a complex set of rules or syntax (see, e.g., Riemann, 1900 or Lerdahl and Jackendoff, 1983). Here, however, we find some important differences. First, in contrast to speech, the rules of musical syntax apply to two dimensions: time (melody) and pitch (harmony). Second, the structural rules in music seem to be much less rigid than syntactic and even prosodic rules in speech. In the domain of semantics, the differences between music and language are even larger. Although music is certainly capable of conveying semantic information (see, e.g., Koelsch *et al.*, 2004), this is not its major purpose. In music, the structure can stand for its own, while in speech it is primarily a framework for the relations between meaningful symbols (words). Another important difference between music and speech concerns the production. While virtually the entire population is in command of at least their mother

tongue, which is learned in the first few years of life, the production of music is limited to a smaller subgroup. In many cases, quite specialized motor skills have to be acquired by years of training.

In the following paragraphs, we present evidence that the processing of musical structure is based on similar brain processes as the processing of the comparable aspects in speech. Furthermore, we deal with the production side of music, in particular with the integration of perception and production mechanisms in the brain.

A. THE PROCESSING OF MUSICAL SYNTAX

A number of ERP studies have been conducted in order to isolate neural correlates for the processing of musical syntax and to compare them to those found with syntax in language. Patel *et al.* (1998) demonstrated that late positive components (P600) were equally elicited by grammatical incongruities in language and harmonic incongruities in music. This ERP component is believed to reflect general knowledge-based integration processes (Friederici, 1995, 1998). Additionally, earlier right lateralized negative components were found in relation to structural expectancy violations around 200 ms (*early right anterior negativity* [ERAN]) (Koelsch *et al.*, 2000) and around 350 ms (*right anterior temporal negativity* [RATN]) (Patel *et al.*, 1998). These components are taken to be specific for music, although similar waves have been found in the domain of speech, like the left anterior negativity (LAN) (Coulson *et al.*, 1998; Kutas and Hillyard, 1983; Rösler *et al.*, 1993) and the early left anterior negativity (ELAN) (Friederici *et al.*, 1993). The latter is elicited especially by word category violations and is taken to reflect initial parsing of a syntactic structure (Friederici, 1995, 1998). Its MEG counterpart has been shown to (partially) originate in left and (to a somewhat lesser extent) right inferior frontal cortices (including Broca's area and its right hemisphere homologue) by magnetic field tomography (Gross *et al.*, 1998), BSCD mapping (Knösche *et al.*, 1999), and current dipole fitting (Friederici *et al.*, 2000a). It would be interesting to know, whether the early syntax-related negativities in music also originate in Broca's area or its right hemisphere homologue. This would cast new light onto the specificity of this brain region, which is well-known for language processing.

For this purpose, sequences of five in-key chords (cadences) were presented to a group of non-musicians, while MEG (148-channel whole-head) was recorded (Maess *et al.*, 2001). In 25% of the sequences, the third chord, in another 25% the fifth chord was replaced by a harmonically unexpected one (Neapolitan). See Fig. 8 for examples. Note that in the fifth position, the degree of unexpectedness is much higher than in the third position. The measured magnetic fields, triggered to the onsets of the third and fifth chords, revealed a difference between Neapolitan

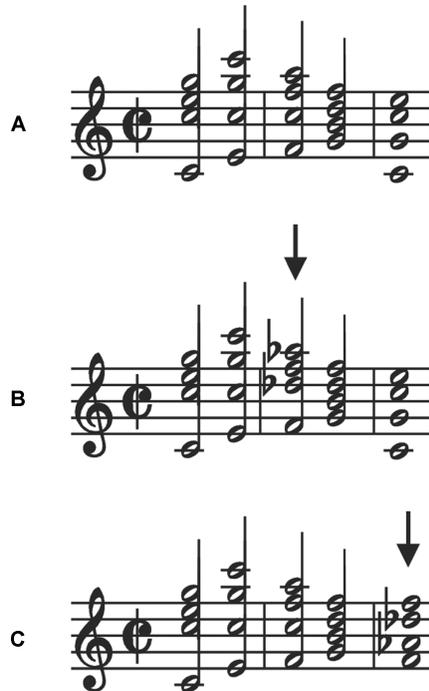


FIG. 8. Example for sequences consisting of in-key chords only (a) and containing one harmonically unexpected (Neapolitan) chord at third (b) and fifth (c) position (adapted from Maess *et al.*, [2001], Fig. 1, with permission of Nature Publishing Group).

and in-key chords. This difference wave was superimposed to the P2m (magnetic counterpart of the P2 in ERP) and had a reversed polarity as well as a more anterior field pattern compared to that component. Its time course was very similar to the one of the ERAN (Koelsch *et al.*, 2000). It was taken to be the magnetic equivalent of the ERAN and, therefore, called ERANm. The ERANm was much larger for the fifth chord (where expectancy was stronger) than for the third one. Dipole modeling localized the generators of the P2m in Heschl's gyrus in each hemisphere, while bilateral activity in the inferior frontal gyrus (presumably A44) was found to underlie the ERANm (Fig. 9).

The described results suggest that the ELAN in speech and the ERAN in music processing are related processes. Both components are elicited by structural expectation violations, they occur between 100 and 200 ms after the onset of the incongruent item, and they localize in or near Broca's area and its right hemisphere homologue. Their hemispherical weighting differs considerably between individuals and seems to reflect interindividual differences in brain organization, although there is a slight overweight of left lateralization for

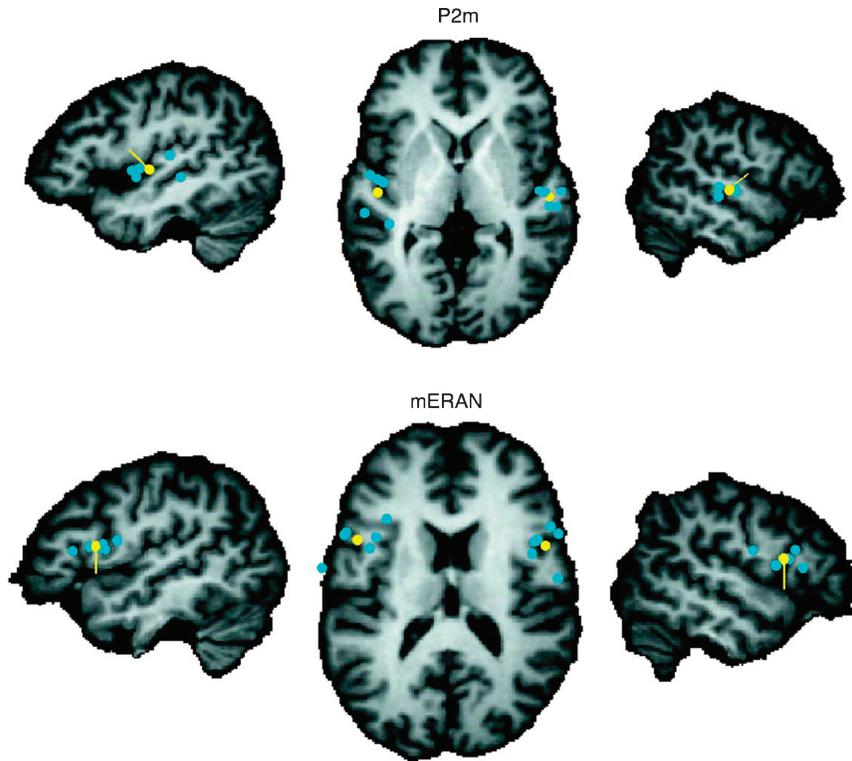


FIG. 9. Source localization results for the P2m and the ERANm. Yellow poles indicate the grand average dipole positions, blue disks represent the single subject solutions (modified from Maess *et al.*, 2001; with permission) (adapted from Maess *et al.*, [2001], Fig. 1, with permission of Nature Publishing Group).

the ELAN and of right lateralization for the ERAN. The conclusion is well supported by an fMRI study using similar musical structure (Koelsch *et al.*, 2002).

Hence, the role of Broca's area and its homologue in fast and automated syntax parsing appears to be less language specific than previously thought.

B. PHRASING IN MUSIC

Both in the domains of music and speech, sequential auditory information must be segmented into phrases in order to allow effective information processing by the human brain. In speech, this segmentation or phrasing is strongly aided by prosodic information, such as rhythm, pauses, accents, amplitude, and pitch variations. In particular, the boundaries between phrases are marked by pauses,

lengthening of the last syllable prior to the boundary and changes (resetting) in the F0 contour. Apart from evidence from behavioral research (Stirling and Wales, 1996; Warren *et al.*, 1995), neuropsychological correlates for the perception and processing of prosodic phrase boundaries by the human brain have also been reported (Steinhauer, 2003; Steinhauer *et al.*, 1999). It was found that each phrase boundary was marked by a positive shift in the ERP of a few hundred milliseconds duration and a centro-parietal scalp distribution. This component was called the *closure positive shift* (CPS), as it was taken to reflect the process of closing the prior phrase. Follow-up studies demonstrated that it was indeed related to *prosodic* information, as it could also be elicited by speech that had been stripped of syntactic or semantic information (Pannekamp *et al.*, 2005; Steinhauer and Friederici, 2001;). It was further shown that the CPS does not necessarily follow each prosodic phrase boundary during speech perception, but that it can shift as a function of the focus structure (Hruska and Alter, 2004; Toepel and Alter, 2003). It does not appear at the prosodic phrase boundary, if the focus is shifted to a particular sub-phrase in the sentence (e.g., when the sentence is preceded by a question). In these cases, the CPS only appears after the relevant (questioned) information is available. Hence, the CPS seems to reflect closure processes that can be the consequence of the perception of prosodic phrase boundaries or focused information, rather than the perception of these boundaries itself.

Phrasing in music bears some striking similarities to the same phenomenon in speech. Its purpose, as in speech, is to aid the perception of structural information (syntax). As in speech, phrase boundaries are marked by local acoustic cues, such as pause insertion and pre-final lengthening, and by more global structural traits, such as melodic contour³. A number of behavioral studies have demonstrated the importance of phrasing for the perception of music by humans (Chiappe and Schmuckler, 1997; Dowling, 1973; Stoffer, 1985; Tan *et al.*, 1981; Wilson *et al.*, 1999). However, so far there was no report of a direct neural correlate, similar to the CPS in speech. therefore, we designed a combined MEG (148 channels whole-head) and EEG (32 channels) study, in order to identify such an electrophysiological marker and to establish its topological and morphological properties as well as its underlying generator structure (Knösche *et al.*, 2005). Musically trained subjects listened to short melodic fragments, each of which was presented in two versions: one clearly divided into two phrases and one consisting of only one phrase (being otherwise identical) (see Fig. 10 for an example). They had to perform a task that was unrelated to the purpose of the study (harmonic decision). EEG and MEG were recorded in separate sessions. In EEG, a centro-parietally distributed positive ERP component, peaking some 550 ms after the offset of the phrase boundary, was found (Fig. 11). The topographic distribution

³For a thorough account of phrase boundary markers in music, see Riemann (1900).



FIG. 10. Example for pair of phrased and unphrased melodies. The unphrased version is constructed from the phrased one by filling the pause at the phrase boundary by one note.

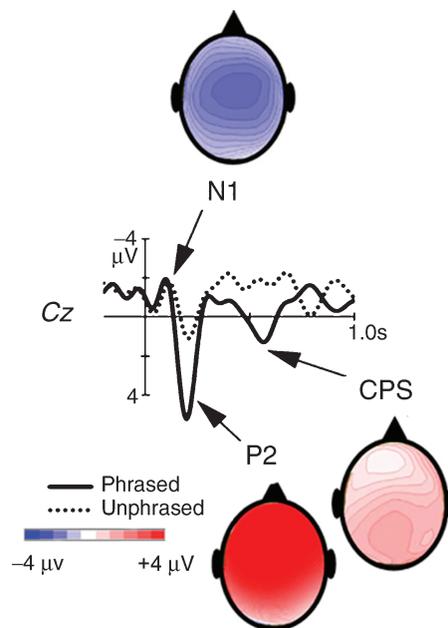


FIG. 11. ERP response to the offset of the phrase boundary: time course of electrode Cz and topographic maps at the peaks of N1, P2, and CPS. The maps refer to the phrased condition.

proved that this component was not created by the superposition of some primary components (e.g., P200) related to the following tones. Moreover, it is very similar to the distribution of the CPS found in speech (Steinhauer *et al.*, 1999). In MEG, we also find a phrase boundary-related component in a similar time range. Its topography is distinct from the respective primary components

(Fig. 12). Its morphology differs from the ERP component, indicating independent information in MEG, as compared to EEG. Subsequent ERP studies (Neuhaus *et al.*, in press) showed that the amplitudes of both components are determined by at least two phrase boundary markers: the length of the inserted pause and the length of the last tone before the pause (prefinal lengthening).

The properties of the ERP components bore some similarity to the CPS found by Steinhauer *et al.* (1999) in relation to the perception of auditory speech, as it was a positive deflection with a centro-parietal maximum without lateralization. On the other hand, the timing was not so perfectly matched between the music and speech-related CPS components. Pannekamp *et al.* (2005) reported data from similar speech material as well as linguistically impoverished speech,

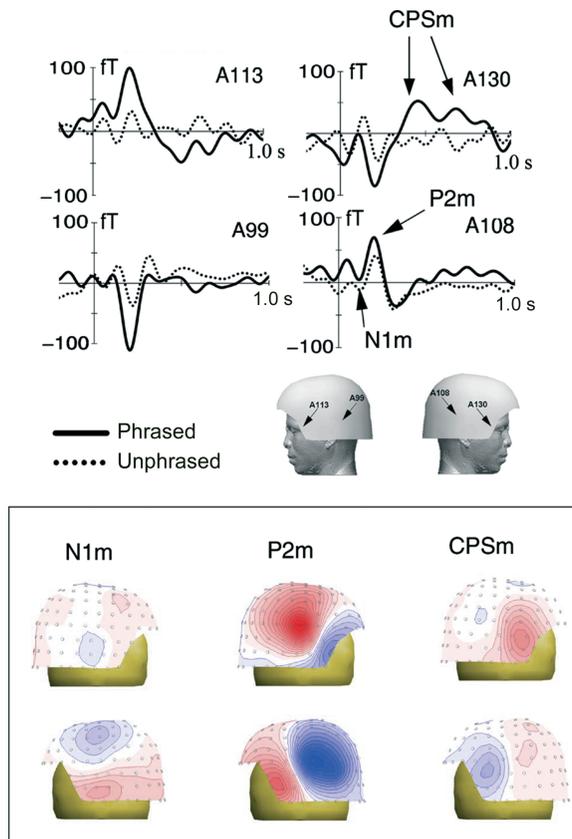


FIG. 12. MEG response to the offset of the phrase boundary: Time courses for selected sensors and topographic maps at the peaks of N1m, P2m, and CPSm. The maps refer to the phrased condition.

where the CPS became significant between 150 and 700 ms after the *onset* of the first phrase boundary. The shorter latency in speech may be due to a relatively larger influence of early parameter changes signaling the end of a phrase, such as lengthened prefinal syllables and changed F0 contours. Nonetheless, we believe that the ERP response to musical phrase boundaries is similar in nature to the CPS in speech and, therefore, we call it *music CPS* and its MEG counterpart is termed *music CPSm*.

In order to gain some information on the underlying neural networks of CPS and CPSm, we applied the multiple signal classification (MUSIC) method (Mosher *et al.*, 1992). This method maps the projection angle of the predicted EEG/MEG due to a probe dipole onto the noise subspace of the data (which is orthogonal to the signal subspace). It is not an inverse method in the strict sense, but a so-called scanning method, since it does not employ a unified model of the activity underlying the measured data (Vrba and Robinson, 2001). It produces maps throughout the brain, each voxel value representing a certain plausibility measure that the respective small brain region might have made a contribution to the measured data. The resulting maps are depicted in Fig. 13. It turns out that both CPS and CPSm are compatible with generators near the posterior and anterior cingulate cortex (PCC, ACC), the subcallosal/medial orbitofrontal area (OFC), the retrosplenial cortex (RC), and the posterior parahippocampal areas (PHC). It should be emphasized that these areas are possible, not necessary, contributors. Most of them belong to the limbic system and are known to be involved in processes concerning attention (ACC), episodic memory retrieval (PCC), and encoding (PHC). One possible interpretation would be that CPS and CPSm reflect the transition from one cognitive period to the next one,

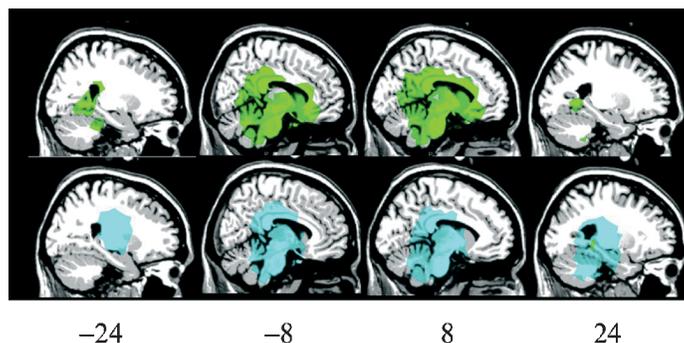


FIG. 13. Results of the source localization based on MEG (400...700 ms after pause offset; upper row, green) and EEG (500...600 ms after pause offset; lower row, turquoise) data. The colored areas show a correlation between the local source space and the signal subspace of less than 7 degrees. The displayed slices are arranged symmetrically around the medial plan from left to right, with a distance of 16 mm. The denoted slice positions are in millimeters on the Talairach x axis.

involving the recognition of the pattern that underlies the previous phrase (phrase closure), storing it as a unified entity (chunking), and redirecting the attention towards the new input (expectation of new phrase). The involvement of cingulate cortex is compatible with recent fMRI findings on prosodic processing of speech (Meyer *et al.*, 2004).

Hence, it seems that the music CPS (and CPSm) does not directly reflect the detection of the phrase boundary as such, but rather memory and attention-related processes accompanying the transition from one cognitive period to the next one. This is compatible with studies showing that the CPS in speech is influenced by the focus structure (Hruska and Alter, 2004; Toepel and Alter, 2003) and the syntactic content of the stimulus material (Pannekamp *et al.*, 2005).

An important question concerns the role of musical expertise. In the speech experiments (e.g., Steinhauer *et al.*, 1999; Steinhauer and Friederici, 2001; Pannekamp *et al.*, 2005), the subjects had expert *implicit*, but little *explicit* knowledge on the presented stimulus material (as they were native speakers, but no linguists). The exact counterpart for this situation in the music domain would involve musicians without formal musical training and knowledge. Since these are not easily recruited, we used musicians with musical training. Therefore, it is interesting to investigate, if the observed brain responses remain the same, if non-musicians are investigated. A recent study addressing this issue (Neuhaus *et al.*, in press) revealed clear differences between the brain responses of musicians and non-musicians (Fig. 14). Instead of a CPS, non-musicians exhibited an early negativity and a less pronounced CPSm, revealing different perceptual strategies for both subject groups. One possible interpretation of this difference would assume that, while musicians process musical phrases similar to language in a structured manner, non-musicians primarily detect the discontinuity in the melodic input. However it remains to be investigated, how this difference holds e.g., and difficult task conditions.

C. SENSORY MOTOR COUPLING

The acquisition of skills is a crucial process in both human development and recovery after pathological conditions. Many skills involve a high degree of automation (e.g. speech, writing, musical performance, or car driving). Such skills and especially the learning thereof are characterized by a close interaction between perceptive concepts and motor programs in a feedback loop. For example, if a pianist is learning a new piece, he/she listens to the sound of her/his own playing and adjusts his/her finger movements accordingly. Another example is the re-acquisition of motor or language skills after a stroke or other cerebral injury. Here, sensory feedback facilitates cortical reorganization and hence the

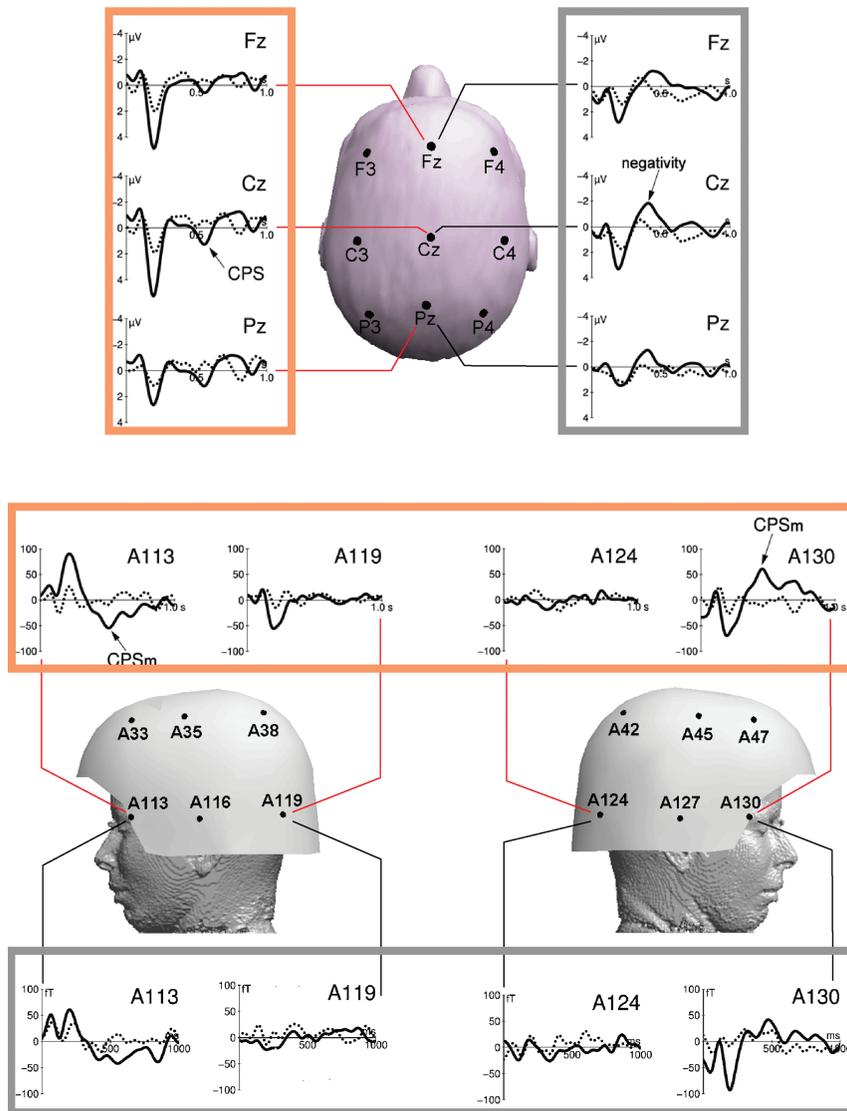


FIG. 14. ERP and MEG responses to phrase boundaries in musicians (red) and non-musicians (black).

re-acquisition of the lost skills. This suggests a very close interaction between the cortical representations of perception and motor programs. The neural basis for such a close interaction between motor actions and the visual observation of the same actions has been identified as the *mirror neuron system* in both non-human

and human primates by Rizzolatti and colleagues (Fogassi *et al.*, 1998; Grafton *et al.*, 1996; Iacoboni *et al.*, 1999, 2001; Rizzolatti *et al.*, 1996a,b).

Musicians need to be in command of a quite extensive repertoire of automated complex sequences of movement. These sequences of movements are naturally very closely related to the produced music, and hence to the respective acoustic percepts. This suggests a strong coupling between the perception and the production of music. Bangert *et al.* (1999) used EEG to investigate this matter in piano-learners. They found that the spatial EEG patterns become more and more similar between the perception and production tasks when the subjects proceed in their learning. This suggested some sort of unified network for instrumental music, which becomes activated in both listening and performing situations. In order to find out which parts of the motor system are *involuntarily* activated by the mere listening to well-trained pieces of music, an MEG study was designed (Haueisen and Knösche, 2001). For voluntary activation of the motor system without actual movement (motor imagery), the involvement of primary motor cortex (M1), supplementary motor area (SMA), premotor cortex (PMC), and anterior cingulate cortex (ACC) has been demonstrated using fMRI (Lotze *et al.*, 1999; Porro *et al.*, 1996; Roth *et al.*, 1996), PET (Stefan *et al.*, 1995), EEG (Beisteiner *et al.*, 1995; McFarland *et al.*, 2000; Pfurtscheller and Neuper, 1997; Pfurtscheller *et al.*, 1999), and MEG (Schnitzler *et al.*, 1997).

As a test case, the perception of piano music was chosen. According to our hypothesis, a person listening to a well-known piece of music (played on a piano) should exhibit involuntary motor cortex activation, if he/she is trained to play this piece on a piano. In contrast, another person who is equally familiar with the musical piece as such, but lacks the playing ability, would not show such an effect. In order to test this hypothesis, two groups of subjects were recruited, which were comparable in their active musical experience, gender, and age. One of them consisted of chorus singers with no piano experience, the other one of pianists. The subject listened to single-voice piano pieces, which were well known to all subjects and well trained by the pianists. They had to listen carefully and perform a tonality judgment task. It was made sure that no actual movements were performed (monitored by EMG) and that no intentional motor imagery was taking place (ensured by using the dominant hand for the task response and checked by subsequent questioning). MEG was measured (148 channels whole-head) and projected onto the brain surface using the BSCD method (Knösche *et al.*, 1996).

While listening to the music, pianists showed significantly increased activity over the hand area of the left primary sensory-motor cortex, as compared to non-pianists. Moreover, when sorting the responses to notes that are normally played by certain fingers (as judged by a musical expert), it turns out that listening to those notes usually played by the thumb cause a motor response about 8 mm more lateral than those notes, which are played by the little finger (Fig. 15). This

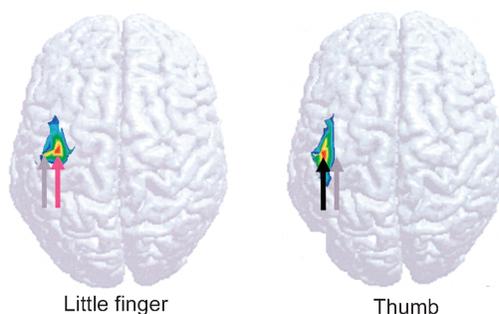


FIG. 15. Brain surface current density (BSCD) mapping of the differences between motor responses to piano notes of pianists and non-pianists. The activity elicited by notes preferably played with the thumb is about 8 mm more lateral than the one elicited by notes played with the little finger.

mimics the homunculus organization of the primary somatosensory and motor cortices and strongly suggests that M1 plays a role here.

The activation of M1 without actual movement has also been found in a number of motor imagery studies using EEG or MEG (Beisteiner *et al.*, 1995; McFarland *et al.*, 2000; Pfurtscheller and Neuper, 1997; Pfurtscheller *et al.*, 1999; Schnitzler *et al.*, 1997). In studies using brain imaging methods like PET or fMRI, the picture is less unanimous. Although a large number of studies report activity in the primary motor cortex (Boecker *et al.*, 2002; Dechent *et al.*, 2004; Grézes and Decety, 2002; Lotze *et al.*, 1999; Malouin *et al.*, 2003; Porro *et al.*, 1996, 2000; Rodriguez *et al.*, 2004; Servos *et al.*, 2002; Stippich *et al.*, 2002; Thobois *et al.*, 2000), many others fail to see such activity related to imagined movements (Binkofski *et al.*, 2000; Decety *et al.*, 1988, 1994; Deiber *et al.*, 1998; Fox *et al.*, 1987; Gerardin *et al.*, 2000; Hanakawa *et al.*, 2003; Ingvar and Philipson, 1977; Meister *et al.*, 2004; Rao *et al.*, 1993; Roland *et al.*, 1980; Samuel *et al.*, 2001; Vingerhoets *et al.*, 2002). Other brain regions directly involved in motor execution have been identified as well in most of the aforementioned studies (e.g., supplementary motor area [SMA], dorsal and ventral premotor cortex, cerebellum, and basal ganglia).

From this literature, it seems clear that large parts of the executive motor system are also activated if the movement is mentally simulated. Apart from such intentional or explicit motor imagery, there are a number of situations where the imagination of movements is implicitly induced by task requirements (see, e.g., Jeannerod and Frak, 1999). These situations include all sorts of mental manipulations of objects, in particular mental rotation paradigms. SMA (Calhoun *et al.*, 2004; Johnston *et al.*, 2004; Richter *et al.*, 2000) and premotor cortex (BA 6) (Cohen *et al.*, 1996; Kosslyn *et al.*, 1998; Lamm *et al.*, 2001; Mecklinger *et al.*, 2002; Richter *et al.*, 2000; Vingerhoets *et al.*, 2002; Wraga *et al.*, 2003) are found

active in these processes. Primary motor cortex involvement is reported by Kosslyn and colleagues (Kosslyn *et al.*, 1998, 2001; Wraga *et al.*, 2003) for such cases, where the task involves pictures of hands or transformations as a consequence of manual activity.

However, not only the intentional or task-induced *imagination* of movements, but also the mere observation of biological movements and actions as well as tools or other graspable objects can activate areas of the motor system (Decety *et al.*, 1997; Grézes and Decety, 2002; Grafton *et al.*, 1997), including primary motor cortex (Hari *et al.*, 1998; Nakamura *et al.*, 2004). Motor areas form part of the human mirror neuron system (Rizzolatti *et al.*, 1996a).

From our results, it seems that auditory input, which is the result of actions the subjects are able to perform, can also activate such a mirror neuron system that includes primary motor cortex. In a recent fMRI study (Meister *et al.*, 2004), no such M1 activity was found when the subjects were reading the musical score and imagine playing a piano. This could mean that the perception of the auditory feedback of a well-trained action works more effective on the motor system than the imagination of the same action without sound. On the other hand, the lack of M1 activity in the study of Meister *et al.* (2004) is just in line with the general disagreement on the involvement of M1 in the motor imagery literature. One possible reason might lie in a special weakness and brevity of this activity (in the case of imagery, in contrast to actual action), which makes it elusive for the brain imaging techniques.

Another question that might arise is why we did not find any other part of the motor system, which is usually very reliably found in all of the aforementioned paradigms, in particular SMA and premotor cortex (BA 6). Possibly, the high level of skill of the subjects rendered the motor programming processes quite automated and, therefore, required little recruitment of the higher motor areas. In the case of SMA, the explanation might also be an anatomical one. It is striking that most EEG and MEG studies on motor imagery do not report activity from this area (e.g., Beisteiner *et al.*, 1995; Schnitzler *et al.*, 1997; Pfurtscheller *et al.*, 1999; McFarland *et al.*, 2000). Several studies in primates and humans have revealed that even unilateral movements cause bilateral activity in the SMA, which would lead to mutual cancellation of the dipolar activity in the medial walls of the frontal lobes and, therefore, explain the failure of MEG to detect the activity (Cheyne and Weinberg, 1989; Lang *et al.*, 1991).

IV. Perception of Gestures

Undisputedly, verbal communication plays the prime role in human interaction. However, there are a number of very important non-verbal means of communication, including behavior, facial expression, eye gaze, and gestures. The most

basic way to obtain information about others' intentions is the observation of their actions. For the recognition of actions, it is postulated that the observed action has to be mapped onto the observer's internal representations of the same action. Such interaction between sensory and motor mechanisms (see also the previous subsection) is supported by the *mirror neuron system*, which has first been identified in monkeys (Fogassi *et al.*, 1998; Rizzolatti *et al.*, 1996a). In humans, this system seems to comprise of a fronto-temporo-parietal network, most consistently including the inferior frontal and parietal cortices as well as the superior temporal sulcus (STS), but also involving parts of the motor system, such as the primary motor cortex, premotor cortex, and preSMA (Decety *et al.*, 1997; Grafton *et al.*, 1996; Hari *et al.*, 1998; Iacoboni *et al.*, 1999, 2001; Nishitani and Hari, 2000; Rizzolatti *et al.*, 1996b).

Other aspects of non-verbal communication are more symbolic in nature. Eye gaze, facial expression, certain types of behavior, and gestures convey information on a person's attitude towards something or somebody as well as an indication to their state of mind. Such signals are thought to be processed in the *social recognition system* (Allison *et al.*, 2000). This system seems to incorporate the superior temporal sulcus, the amygdala, the orbito-frontal cortex, and the right somatosensory cortex (Adolphs, 1999; Allison *et al.*, 2000; Grossmann *et al.*, 2000).

Symbolic hand signs (emblematic gestures) (McNeill, 1992) are a unique type of gesture, because they can transmit intentions independent of speech, like "okay," "good," or "bad." However, the underlying brain mechanisms were undefined. The recognition of hand signs seems to involve (all) aspects of object/shape recognition, action recognition, and the recognition of social signals. Thus, the different neuronal networks underlying these processes should work together to achieve unified perception/recognition. We designed an MEG study (Nakamura *et al.*, 2004) to reveal orchestrated multiple brain activity across different neuronal systems, including *object recognition system* (ventral occipito-temporal; see, e.g., Ungerleider and Haxby, 1994), *mirror neuron system*, and *social recognition system* during the recognition of hand signs.

In order to disentangle the contributions of the three subsystems of hand sign recognition, two different experimental conditions were presented to the subjects. In the *meaning task* condition, subjects had to discriminate meaningful hand signs (HM+) from meaningless hand postures (HM-)⁴, which were matched in shape as much as possible (see Fig. 16 for some examples). In a second condition, the *categorization task*, the meaningful hand signs (HC) had to be distinguished from faces.

MEG was measured using a 148-channel whole-head system (MAGNES WHS 2500, 4D-Neuroimaging Inc., San Diego, CA). The BSCD maps were computed using head models, which were based on individual MRI data. The

⁴The 11 pairs of meaningful and meaningless hand postures were selected from 66 different images of hand postures based on the rating of 47 German native speakers.

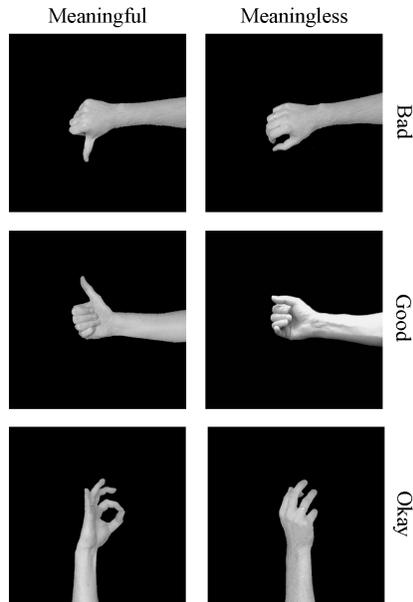


FIG. 16. Some examples out of the 11 different pairs of meaningful hand signs and shape-matched meaningless hand postures (adapted from Nakamura *et al.*, [2004], Fig. 1 and 4, with permission of Elsevier).

results were transformed to Talairach space and then treated with factor analysis based on principal components (PCA) (Maess *et al.*, 2002)⁵. The PCA could extract 10 spatio-temporally separable factors from the CSD data, which showed physiologically plausible spatial distributions. Using the topographical information from these PCA factors, we created 13 regions of interest (ROIs) by thresholding⁶. Then, the time courses for the three conditions (HM+, HM-, and HC) in each of the 13 ROIs were averaged and statistically tested⁷.

The results are summarized in Fig. 17, which depicts a possible model of the information processing during hand sign recognition. The model consists of processes in the primary visual, the mirror neuron, the social recognition, and the object recognition system. Regions which showed significantly stronger

⁵Source position formed one dimension of the decomposed matrix, while time instant (between -100 and 660 ms), conditions, and subjects were arranged along the other dimension. Following cutoff after ten factors, VARIMAX rotation was applied.

⁶In some cases, the factor map consisted of more than two distinct centers of activity, leading to the extraction of several ROIs.

⁷Condition effects were tested using ANOVAs (time and condition) with subsequent post hoc testing. Bonferroni correction for multiple testing (one test per ROI) was applied.

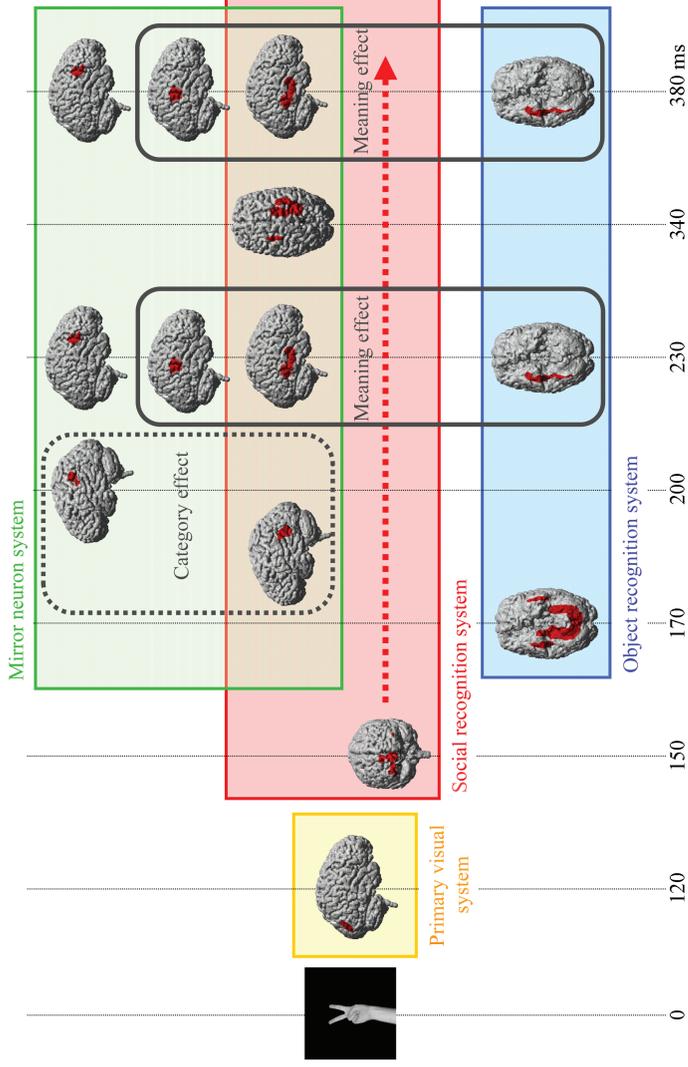


Fig. 17. Schematic diagram of possible brain mechanisms underlying the recognition of hand signs. The active brain areas are arranged according to their peak latency and grouped into different neuronal systems (modified from Nakamura *et al.*, 2004).

activation in the HM+ condition (HM+ > HM- and/or HC: meaning effect) at a given time window and those that showed significant activation in the HC condition (category effect) are marked in Fig. 17. After initial processing in the primary visual system peaking around 120 ms, different aspects of information are processed in parallel in anatomically distinct brain areas (about 170 to 200 ms; left inferior parietal cortex and posterior STS). Thereafter, distinct brain areas are simultaneously activated in the HM+ condition (230 ms) suggesting cooperation across different neuronal systems during the assessment of the meaning of hand signs. At around 340 ms, the somatomotor cortices are maximally active, which probably reflect mapping processes onto the internal representations (Decety and Grézes, 1999; Adolphs, 1999). Finally, different brain regions are again activated in concert (370 to 380 ms). The right inferior prefrontal region was continuously activated after the latency about 150 ms suggesting top-down and/or supervisory processes.

It is striking that there is a marked hemispheric asymmetry suggesting functional dissociation between the right and left hemisphere in hand sign recognition. The left inferior parietal and STS regions are activated during categorization, whereas the right inferior parietal and STS as well as inferior temporal regions are activated during assessing the meaningfulness of the hand signs. In most studies investigating action perception, a left hemispheric predominance was observed (Decety *et al.*, 1997; Grafton *et al.*, 1996) similar to the language processing. On the other hand, it is suggested that emotional and social recognition, in particular the recognition of facial expressions, are processed mainly in the right hemisphere (Adolphs *et al.*, 1999; Nakamura *et al.*, 1999). Hence, it seems that the meaning of hand signs are interpreted as a kind of social signals, similar to face expressions, and processed in the right hemisphere.

V. Conclusions

We have presented a number of studies that demonstrate the ability of MEG to identify brain areas involved in the perception of speech, music, and hand signs. These generators are located in all cortical lobes as well as in subcortical areas. Most importantly, however, MEG could attribute a precise timing to the localized brain processes (e.g., Maess *et al.*, 2002). This enabled us to construct even complex spatio-temporal models of the brain processes underlying cognitive function (Knösche *et al.*, 2002; Nakamura *et al.*, 2004).

The diversity of the employed modeling techniques highlights the fact that MEG- (and EEG-) based source reconstruction is not a routine business. The choice of forward and inverse methods has to be made anew in each case,

carefully considering the nature and quality of the data, available supporting information (e.g., MRI), and the hypotheses on the possible outcome.

In the future, considerable improvements are expected, if additional information on the brain activity or the surrounding head tissue is made available. This includes, for example, the use of anisotropic tissue properties by the FEM method, the exploitation of trial-by-trial variability (single-trial analysis), introduction of additional functional knowledge (e.g., by fMRI), or anatomical details (e.g., connectivity information from tractography based on diffusion weighted MRI). The further integration of the various source of functional and anatomical information might be the way towards more realistic and instructive models of the intracranial processes. This, however, requires an even more detailed understanding of the underlying mechanisms for each of the techniques.

References

- Adolphs, R. (1999). Social cognition and the human brain. *Trends Cogn. Sci.* **3**, 469–479.
- Alho, K., Tervaniemi, M., Huottilainen, M., Lavikainen, J., Tiitinen, H., Ilmoniemi, R. J., Knuutila, J., and Näätänen, R. (1996). Processing of complex sounds in the human auditory cortex as revealed by magnetic brain responses. *Psychophysiology* **33**, 369–375.
- Alho, K., Connolly, J. F., Cheour, M., Lehtokoski, A., Huottilainen, M., Virtanen, J., Aulanko, R., and Ilmoniemi, R. J. (1998). Hemispheric lateralization in preattentive processing of speech sounds. *Neurosci. Lett.* **258**, 9–12.
- Allison, T., Puce, A., and McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends Cogn. Sci.* **4**, 267–278.
- Bangert, M., Parlitz, D., and Altenmüller, E. (1999). An interface for complex auditory-sensorimotor integration: Where the pianist's cortex maps perception to action. *Neuroimage* **9**(part 2), S419.
- Broca, P. (1861). Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie (parte de la parole). *Bull. Soc. Anat. Paris* **6**, 330–357.
- Beisteiner, R., Hollinger, P., Lindinger, G., Lang, W., and Berthoz, A. (1995). Mental representation of movements. Brain potentials associated with imagination of hand movements. *Exp. Brain Res.* **78**, 604–612.
- Belin, P., Zatorre, R., and Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Cogn. Brain Res.* **13**, 17–26.
- Binkofski, F., Amunts, K., Stephan, K. M., Posse, S., Schormann, T., Freund, H. J., Zilles, K., and Seitz, R. J. (2000). Broca's region subserves imagery of motion: A combined cytoarchitectonic and fMRI study. *Human Brain Mapping* **11**, 273–285.
- Boecker, H., Ceballos-Baumann, A. O., Bartenstein, P., Dagher, A., Forster, K., Haslinger, B., Brooks, D. J., Schwaiger, M., and Conrad, B. (2002). A H₂¹⁵O positron emission tomography study on mental imagery of movement sequences—the effect of modulating sequence length and direction. *NeuroImage* **17**, 999–1009.
- Calhoun, V. D., Altschul, D., McGinty, V., Shih, R., Scott, D., Sears, E., and Pearlson, G. D. (2004). Alcohol intoxication effects on visual perception: An fMRI study. *Human Brain Mapping* **21**, 15–25.
- Caplan, D., Alpert, N., and Waters, G. (1998). Effects of syntactic structure and proposition number on patterns of regional cerebral blood flow. *J. Cogn. Neurosci.* **10**, 541–552.

- Cheyne, D., and Weinberg, H. (1989). Neuromagnetic fields accompanying unilateral finger movements: Pre-movement and movement-evoked fields. *J. Cogn. Neurosci.* **11**, 491–501.
- Chiappe, P., and Schmuckler, M. A. (1997). Phrasing influences the recognition of melodies. *Psychon. B. Rev.* **4**, 254–259.
- Coulson, S., King, J., and Kutas, M. (1998). Expect the unexpected: Event-related brain responses of morphosyntactic violations. *Lan. Cog. Proc.* **13**, 21–58.
- Csepe, V., Pantev, C., Hoke, M., Hampson, S., and Ross, B. (1992). Evoked magnetic responses of the human auditory cortex to minor pitch changes—localization of the mismatch field. *Electroenceph. Clin. Neurophysiol.* **84**, 538–548.
- Damian, M. F., Vigliocco, G., and Levelt, W. J. M. (2001). Effects of semantic context in the naming of pictures and words. *Cognition* **81**, B77–B86.
- Dechent, P., Merboldt, K., and Frahm, J. (2004). Is the human primary motor cortex involved in motor imagery? *Cogn. Brain Res.* **19**, 138–144.
- Decety, J., Phillipson, B., and Ingvar, D. H. (1988). rCBF landscapes during motor performance and motor ideation of graphic gesture. *Arch. Psychiat. Neurol. Sci.* **238**, 33–38.
- Decety, J., Grèzes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F., and Fazio, F. (1997). Brain activity during observation of actions—influence of action content and subject's strategy. *Brain* **120**, 1763–1777.
- Decety, J., Perani, D., Jeannerod, M., Bettinardi, V., Tadary, B., Woods, R., Mazziotta, J. C., and Fazio, F. (1994). Mapping motor representations with PET. *Nature* **371**, 600–602.
- Decety, J., and Grezes, J. (1999). Neural mechanisms subserving the perception of human actions. *Trends Cogn. Sci.* **3**, 172–178.
- Deiber, M., Ibañez, V., Honda, M., Sadato, N., Raman, R., and Hallet, M. (1998). Cerebral processes related to visuomotor imagery and generation of simple finger movements studied with positron emission tomography. *NeuroImage* **7**, 73–85.
- Démonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., Rascol, A., and Frackowiak, R. S. J. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain* **115**, 1753–1768.
- Diesch, E., and Luce, T. (1997). Magnetic mismatch fields elicited by vowels and consonants. *Exp. Brain Res.* **116**, 139–152.
- Dowling, W. J. (1973). Rhythmic groups and subjective chunks in memory for melodies. *Percept. Psychophys.* **14**, 37–40.
- Dronkers, N. E., Redfern, B. B., and Ludy, C. A. (1995). Lesion localization in chronic Wernicke's aphasia. *Brain Lang.* **51**, 62–65.
- Fox, P. T., Pardo, J. V., Petersen, S. E., and Raichle, M. E. (1987). Supplementary motor and premotor responses to actual and imagined hand movements with positron emission tomography. *Soc. Neurosci. Abstr.* **13**, 1433.
- Fogassi, L., Gallese, V., Fadiga, L., and Rizzolatti, G. (1998). Neurons responding to the sight of goal directed hand/arm actions in the parietal area PF (7b) of the macaque monkey. *Soc. Neurosci. Abstr.* **24**, 257.
- Friederici, A. D., Pfeifer, E., and Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cogn. Brain Res.* **1**, 183–192.
- Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain Lang.* **50**, 259–281.
- Friederici, A. D., ed. (1998). "Language Comprehension: A Biological Perspective." Springer, Berlin.
- Friederici, A. D., Hahne, A., and von Cramon, D. Y. (1998). First-pass versus second-pass parsing processes in a Wernicke's and a Broca's aphasic: Electro-physiological evidence for a double dissociation. *Brain Lang.* **62**, 311–341.

- Friederici, A. D., von Cramon, D. Y., and Kotz, S. A. (1999). Language related brain potentials in patients with cortical and subcortical left hemisphere lesions. *Brain* **122**, 1033–1047.
- Friederici, A. D., Wang, Y., Herrmann, C. S., Maess, B., and Oertel, U. (2000a). Localization of early syntactic processes in frontal and temporal cortical areas: A magnetoencephalographic study. *Human Brain Mapping* **11**, 1–10.
- Friederici, A. D., Meyer, M., and von Cramon, D. Y. (2000b). Auditory language comprehension: An event-related fMRI study on the processing of syntactic and lexical information. *Brain Lang.* **74**, 289–300.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends Cogn. Sci.* **6**, 78–84.
- Friederici, A. D., Rüschemeyer, S., Hahne, A., and Fiebach, C. J. (2003). The role of left anterior frontal and superior temporal cortex in sentence comprehension: Localizing syntactic and semantic processes. *Cerebral Cortex* **13**, 170–177.
- Friederici, A. D., and Kotz, S. A. (2003). The brain basis of syntactic processes: Functional imaging and lesion studies. *Neuroimage* **20**(Suppl 1), S8–S17.
- FrodlBauch, T., Kathmann, N., Moller, H. J., and Hegerl, U. (1997). Dipole localization and test-retest reliability of frequency and duration mismatch negativity generator processes. *Brain Topography* **10**, 3–8.
- Gerardin, E., Sirigu, A., Lehérici, S., Polin, J., Gaymard, B., Marsault, C., Agid, Y., and Le Bihan, D. (2000). Partially overlapping neural networks for real and imagined hand movements. *Cerebral Cortex* **10**, 1093–1104.
- Giard, M. H., Lavikainen, J., Reinikainen, K., Perrin, F., Bertrand, O., Pernier, J., and Näätänen, R. (1995). Separate representations of stimulus frequency, intensity and duration in auditory sensory memory: An event-related potential and dipole-model analysis. *J. Cogn. Neurosci.* **7**, 133–143.
- Grafton, S. T., Arbib, M. A., Fadiga, L., and Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography 2. Observation compared with imagination. *Exp. Brain Res.* **112**, 103–111.
- Grafton, S. T., Fadiga, L., Arbib, M. A., and Rizzolatti, G. (1997). Premotor cortex activation during observation and naming of familiar tools. *NeuroImage* **6**, 231–236.
- Grézes, J., and Decety, J. (2002). Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia* **40**, 212–222.
- Griffith, T. D., and Warren, J. D. (2002). The planum temporale as computational hub. *Trends Neurosci.* **25**, 348–353.
- Gross, J., Ioannides, A. A., Dammers, J., Maess, B., Friederici, A. D., and Müller-Gärtner, H.-J. (1998). Magnetic field tomography analysis of continuous speech. *Brain Topography* **10**, 273–281.
- Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., and Blake, R. (2000). Brain areas involved in perception of biological motion. *J. Cogn. Neurosci.* **12**, 711–720.
- Hahne, A. (1998). “Charakteristika syntaktischer und semantischer Prozesse bei der auditiven Sprachverarbeitung: Evidenz aus ereigniskorrelierten Potentialstudien - MPI Series in Cognitive Neuroscience.” I. Risse, Leipzig.
- Hämäläinen, M. S., and Sarvas, J. (1987). Feasibility of the homogeneous head model in the interpretation of neuromagnetic data. *Phys. Med. Biol.* **32**, 91–97.
- Hämäläinen, M., and Ilmoniemi, R. (1994). Interpreting magnetic fields of the brain: Minimum norm estimates. *Med. Biol. Comp.* **32**, 183.
- Hanakawa, T., Immisch, I., Toma, K., Dimyan, M. A., van Gelderen, P., and Hallett, M. (2003). Functional properties of brain areas associated with motor execution and imagery. *J. Neurophysiol.* **89**, 989–1002.

- Hari, R., Hämäläinen, M., Ilmoniemi, R., Kaukoranta, E., Reinikainen, K., Salminen, J., Alho, K., Näätänen, R., and Sams, M. (1984). Response of the primary auditory cortex to pitch changes in a sequence of tone pips: Neuromagnetic recordings in man. *Neurosci. Lett.* **50**, 127–132.
- Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., and Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: A neuromagnetic study. *PNAS* **95**, 15061–15065.
- Haueisen, J., and Knösche, T. R. (2001). Involuntary motor activation in pianists evoked by music perception. *J. Cogn. Neurosci.* **13**, 786–792.
- Hruska, C., and Alter, K. (2004). How prosody can influence sentence perception. In “Information Structure: Theoretical and Empirical Evidence” (A. Steube, Ed.). Walter de Gruyter, Berlin.
- Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., and Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science* **286**, 2526–2528.
- Iacoboni, M., Koski, L. M., Brass, M., Bekkering, H., Woods, R. P., Dubeau, M. C., Mazziotta, J. C., and Rizzolatti, G. (2001). Reafferent copies of imitated actions in the right superior temporal cortex. *PNAS* **98**, 13995–13999.
- Ingvar, D. H., and Philipson, L. (1977). Distribution of cerebral blood-flow in dominant hemisphere during motor ideation and motor performance. *Annf. Neurol.* **2**, 230–237.
- Ioannides, A. A., Bolton, J. P. R., and Clarke, C. J. S. (1990). Continuous probabilistic solutions to the biomagnetic inverse problem. *Inverse Problems* **6**, 523–542.
- Jaeger, J. J., Lockwood, A.H., Vanvalin, R. D., Kemmerer, D. L., Murphy, B. W., and Wack, D. S. (1998). Sex differences in brain regions activated by grammatical and reading task. *NeuroReport* **9**, 2803–2807.
- Jäncke, L., Wusenberg, T., Scheich, H., and Heinze, H. (2002). Phonetic perception and the temporal cortex. *NeuroImage* **15**, 733–746.
- Jeannerod, M., and Frak, V. (1999). Mental imaging of motor activity in humans. *Curr Opin Neurobiol* **9**, 735–739.
- Johnston, S., Leek, E. C., Atherton, C., Thacker, N., and Jackson, A. (2004). Functional contribution of medial premotor cortex to visuo-spatial transformation in humans. *Neurosci. Lett.* **355**, 209–212.
- Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., and Thulborn, K. R. (1996). Brain activation modulated by sentence comprehension. *Science* **274**, 114–116.
- Knösche, T. R., Praamstra, P., Stegeman, D., and Peters, M. J. (1996). Linear estimation discriminates midline sources and a motor cortex contribution to the readiness potential. *Electroenceph. Clin. Neurophysiol.* **99**, 183–190.
- Knösche, T. R. (1997). “Solutions of the Neuroelectromagnetic Inverse Problem—An Evaluation Study” PhD thesis, University of Twente, The Netherlands.
- Knösche, T. R., Maess, B., and Friederici, A. D. (1999). Processing of syntactic information monitored by brain surface current density mapping based on MEG. *Brain Topography* **12**, 75–87.
- Knösche, T. R., Lattner, S., Maess, B., Schauer, M., and Friederici, A. D. (2002). Early parallel processing of auditory word and voice information. *NeuroImage* **17**, 1493–1503.
- Knösche, T. R., Neuhaus, C., Haueisen, J., Alter, K., Maess, B., Witte, O. W., and Friederici, A. D. (2005). The Perception of Phrasing in Music. *Human Brain Mapping* **24**, 259–273.
- Koelsch, S., Gunter, T. C., Friederici, A. D., and Schroeger, E. (2000). Brain indices of music processing: “non-musicians” are musical. *J. Cogn. Neurosci.* **12**, 520–541.
- Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., and Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nat. Neurosci.* **7**, 302–305.
- Koelsch, S., Gunter, T. C., von Cramon, D. Y., Zysset, S., Lohmann, G., and Friederici, A. D. (2002). Bach speaks: A cortical ‘language network’ serves the processing of music. *NeuroImage*, **17**, 956–966.

- Kosslyn, S. M., Digirolamo, G. J., Thompson, W. L., and Alpert, N. M. (1998). Mental rotation of objects versus hands: Neural mechanisms revealed by positron emission tomography. *Psychophysiology* **35**, 151–161.
- Kosslyn, S. M., Thompson, W. L., Wraga, M., and Alpert, N. A. (2001). Imagining rotation by endogenous versus exogenous forces: Distinct neural mechanisms. *NeuroReport* **12**, 2519–2525.
- Kroll, J. F., and Steward, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *J. Memo. Lang.* **33**, 149–174.
- Kutas, M., and Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflecting semantic incongruity. *Science* **207**, 203–205.
- Kutas, M., and Hillyard, S. A. (1983). Event-related potentials to grammatical errors and semantic anomalies. *Memory and Cognition* **11**, 539–550.
- Lake, D. A., and Bryden, M. P. (1976). Handedness and sex differences in hemispheric asymmetry. *Brain Lang.* **3**, 266–282.
- Lamm, C., Windischberger, C., Leodolter, U., Moser, E., and Bauer, H. (2001). Evidence for premotor cortex activity during dynamic visuospatial imagery from single-trial functional magnetic resonance imaging and event-related slow cortical potentials. *NeuroImage* **14**, 268–283.
- Lang, W., Cheyne, D., Kristeva, R., Beisteiner, R., and Lindinger, G. (1991). Three-dimensional localization of SMA activity preceding voluntary movement. A study of electric and magnetic fields in a patient with infarction of the right supplementary motor area. *Exp. Brain Res.* **87**, 688–695.
- Lattner, S., Maess, B., Wang, Y., Schauer, M., Alter, K., and Friederici, A. D. (2003). Dissociation of human and computer voices in the brain: Evidence for a preattentive gestalt-like perception. *Human Brain Mapping* **20**, 13–21.
- Lerdahl, F., and Jackendoff, R. (1983). “A Generative Theory of Tonal Music.” MIT Press, Cambridge.
- Levelt, W. J. M., Praamstra, P., Meyer, A. S., Helenius, P., and Salmelin, R. (1998). An MEG study of picture naming. *J. Cogn. Neurosci.* **10**, 553–567.
- Levelt, W. J. M., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences* **22**, 1–38.
- Levelt, W. J. M., and Indefrey, P. (2000). The speaking mind/brain: Where do spoken words come from? In “Image, Language, Brain” (A. Maranz, Y. Miyatshita, and W. Y. O’Neil, Eds.), pp. 77–93. MIT Press, Cambridge.
- Lotze, M., Montoya, P., Erb, M., Hulsmann, E., Flor, H., Klose, U., Birbaumer, N., and Grodd, W. (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: An fMRI study. *J. Cogn. Neurosci.* **11**, 491–501.
- Lütkenhöner, B., and Steinsträter, O. (1998). High-precision neuromagnetic study of the functional organization of the human auditory cortex. *Audiology and Neuro-Otology* **3**, 191–213.
- Maess, B., Koelsch, S., Gunter, T. C., and Friederici, A. D. (2001). Musical syntax is processed in Broca’s area: An MEG study. *Nature Neuroscience* **4**, 540–545.
- Maess, B., Friederici, A. D., Damian, M., Meyer, A. S., and Levelt, W. J. M. (2002). Semantic category interference in overt picture naming: Sharpening current density localization by PCA. *J. Cogn. Neurosci.* **14**, 455–462.
- Malouin, F., Richards, C. L., Jackson, P. L., Dumas, F., and Doyon, J. (2003). Brain activations during motor imagery of locomotor-related tasks: A PET study. *Human Brain Mapping* **19**, 47–62.
- Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., Salomon, G., Dehaene, S., Cohen, L., and Mehler, J. (1993). The cortical representation of speech. *J. Cogn. Neurosci.* **5**, 467–479.

- McFarland, D. J., Miner, L. A., Vaughan, T. M., and Wolpaw, J. R. (2000). Mu and beta rhythm topographies during motor imagery and actual movements. *Brain Topography* **12**, 177–186.
- McGlone, J. (1978). Sex differences in functional brain asymmetry. *Cortex* **14**, 122–128.
- McNeill, D. (1992). “Hand and Mind: What Gestures Reveal about Thought.” Univ. of Chicago press, Chicago.
- Mecklinger, A., Gruenewald, C., Besson, M., Magnie, M. N., and von Cramon, D. Y. (2002). Separable neuronal circuitries for manipulable and non-manipulable objects in working memory. *Cerebral Cortex* **12**, 1115–1123.
- Meister, I. G., Krings, T., Foltys, H., Boroojerdi, B., Müller, M., Töpper, R., and Thron, A. (2004). Playing piano in the mind - an fMRI study on music imagery and performance in pianists. *Cogn. Brain Res* **19**, 219–228.
- Meyer, M., Friederici, A. D., and von Cramon, D. Y. (2000). Neurocognition of auditory sentence comprehension: Event-related fMRI reveals sensitivity to syntactic violations and task demands. *Cogn. Brain Res.* **9**, 19–33.
- Meyer, M., Steinhauer, K., Alter, K., Friederici, A. D., and von Cramon, D. Y. (2004). Brain activity varies with modulation of dynamic pitch variance in sentence melody. *Brain Lang.* **89** (Special Issue), 277–289.
- Miller, C. E., and Hendriquez, C. S. (1990). Finite element analysis of bioelectric phenomena. *Crit. Rev. Biomed. Eng.* **18**, 207–233.
- Mosher, J. C., Lewis, P. S., and Leahy, R. M. (1992). Multiple dipole modeling and localization from spatio-temporal MEG data. *IEEE T. Biomed. Eng.* **39**, 541–557.
- Mullenix, J. W., and Pisoni, D. B. (1990). Stimulus variability and processing dependencies in speech perception. *Perception and Psychophysics* **47**, 379–390.
- Näätänen, R., Gaillard, A. W. K., and Mantysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica* **42**, 313–329.
- Näätänen, R., Paavilainen, P., Tiitinen, H., Jiang, D., and Alho, K. (1993). Attention and mismatch negativity. *Psychophysiology* **30**, 436–450.
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., and Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* **385**, 432–434.
- Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., and Winkler, I. (2001). ‘Primitive intelligence’ in the auditory cortex. *Trends in Neuroscience* **24**, 283–288.
- Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology* **38**, 1–21.
- Nakamura, K., Kawashima, R., Ito, K., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Nagumo, S., Kubota, K., Fukuda, H., and Kojima, S. (1999). Activation of the right inferior frontal cortex during assessment of facial emotion. *J. Neurophys.* **82**, 1610–1614.
- Nakamura, A., Maess, B., Knösche, T. R., Gunter, T. C., Bach, P., and Friederici, A. D. (2004). Cooperation of different neuronal systems during hand sign recognition. *NeuroImage* **23**, 25–34.
- Neuhaus, C., Knösche, T. R., and Friederici, A. D. (in press). Effects of musical expertise and boundary markers on phrase perception in music. *J. Cogn. Neurosci.*
- Niedermeyer, E., and Lopes da Silva, F., eds. (1999). “Electroencephalography: Basic Principles, Clinical Applications, and Related Fields.” William and Wilkins, Baltimore.
- Nishitani, N., and Hari, R. (2000). Temporal dynamics of cortical representation for action. *PNAS* **97**, 913–918.
- Pannekamp, A., Toepel, U., Alter, K., Hahne, A., and Friederici, A. D. (2005). Prosody driven sentence processing: An ERP study. *J. Cogn. Neurosci.* **17**, 407–421

- Patel, A. D., Gibson, E., Ratner, J., Besson, M., and Holcomb, P. (1998). Processing syntactic relations in language and music: An event-related potential study. *J. Cogn. Neurosci.* **10**, 717–733.
- Penhune, V. B., Zatorre, R. J., Mac Donald, J. D., and Evans, A. C. (1996). Interhemispheric anatomical differences in human primary auditory cortex: Probabilistic mapping and volume measurement from magnetic resonance scans. *Cerebral Cortex* **6**, 661–672.
- Pfurtscheller, G., and Neuper, C. (1997). Motor imagery activates primary sensorimotor area in humans. *Neurosci. Lett.* **239**, 65–68.
- Pfurtscheller, G., Neuper, C., Ramoser, H., and Muller-Gerking, J. (1999). Visually guided motor imagery activates sensorimotor areas in humans. *Neurosci. Lett.* **269**, 153–156.
- Porro, C. A., Francescato, M. P., Cettolo, V., Diamond, M. E., Baraldi, P., Zuiani, C., Bazzocchi, M., and di Prampero, P. E. (1996). Primary motor and sensory cortex activation during motor performance and motor imagery: A functional magnetic resonance imaging study. *J. Neurosci.* **16**, 7688–7698.
- Porro, A. P., Cettolo, V., Francescato, M. P., and Baraldi, P. (2000). Upsilon lateral involvement of primary motor cortex during motor imagery. *Eur. J. Neurosci.* **12**, 3059–3063.
- Rao, S. M., Binder, J. R., Bandettini, P. A., Hammeke, T. A., Yetkin, F. Z., Jemanowicz, A., Lisk, L. M., Morris, G. L., Mueller, W. M., Estkowski, L. D., Wong, E. C., Haughton, V. M., and Hyde, J. S. (1993). Functional magnetic resonance imaging of complex human movements. *Neurology* **43**, 2311–2318.
- Richter, W., Somorjai, R., Summers, R., Jarmasz, M., Menon, R. S., Gati, J. S., Georgopoulos, A. P., Tegeler, C., Ugurbil, K., and Kim, S. G. (2000). Motor area activity during mental rotation studied by time-resolved single-trial fMRI. *J. Cogn. Neurosci.* **12**, 310–320.
- Riemann, H. (1900). “Vademecum der Phrasierung.” Max-Hesse-Verlag, Leipzig.
- Rinne, T., Alho, K., Alku, P., Holi, M., Sinkkonen, J., Virtanen, J., Bertrand, O., and Näätänen, R. (1999). Analysis of speech sounds is left-hemisphere pre-dominant 150 ms after sound onset. *NeuroReport* **10**, 1113–1117.
- Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996a). Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* **3**, 131–141.
- Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., and Fazio, F. (1996b). Localization of grasp representations in humans by PET.1. Observation versus execution. *Exp. Brain Res.* **111**, 246–252.
- Rodríguez, M., Muñoz, R., González, B., and Sabaté, M. (2004). Hand movement distribution in the motor cortex: The influence of a concurrent task and motor imagery. *NeuroImage* **22**, 1480–1491.
- Roland, P. E., Larsen, B., Larsen, N. A., and Skinhoj, E. (1980). Supplementary motor area and other cortical areas in organization of voluntary movements in man. *J. Neurophysiol.* **43**, 118–136.
- Rösler, F., Friederici, A. D., Pütz, P., and Hahne, A. (1993). Event-related brain potentials while encountering semantic and syntactic constraint violations. *J. Cognitive Neurosci.* **5**, 345–362.
- Roth, M., Decety, J., Raybaudi, M., Massarelli, R., Delon-Martin, C., Segebarth, C., Morand, S., Gemignani, A., Decorps, M., and Jeannerod, M. (1996). Possible involvement of primary motor cortex in mentally simulated movement: A functional magnetic resonance imaging study. *NeuroReport* **7**, 1280–1284.
- Sams, M., Paavilainen, P., Alho, K., and Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroenceph. Clin. Neurophysiol.* **62**, 437–448.
- Samuel, M., Ceballos-Baumann, A. O., Boecker, H., and Brooks, D. J. (2001). Motor imagery in normal subjects and Parkinson’s disease patients: An H₂¹⁵O PET study. *NeuroReport* **12**, 821–828.
- Scherg, M., and Berg, P. (1991). Use of prior knowledge in brain electromagnetic source analysis. *Brain Topography.* **4**, 143–150.
- Schnitzler, A., Salenius, S., Salmelin, R., Jousmaki, V., and Hari, R. (1997). Involvement of primary motor cortex in motor imagery: A neuromagnetic study. *NeuroImage* **6**, 201–208.

- Servos, P., Osu, R., Santi, A., and Kawato, M. (2002). The neural substrates of biological motion perception: An fMRI study. *Cerebral Cortex* **12**, 772–782.
- Shtyrov, Y., Kujala, T., Ahveninen, J., Tervaniemi, M., Alku, P., Ilmoniemi, R. J., and Näätänen, R. (1998). Background acoustic noise and the hemispheric lateralization of speech processing in the human brain: Magnetic mismatch negativity study. *Neurosci. Lett.* **251**, 141–144.
- Shtyrov, Y., Kujala, T., Lyytinen, H., Kujala, J., Ilmoniemi, R. J., and Näätänen, R. (2000). Lateralization of speech processing in brain as indicated by mismatch negativity and dichotic listening. *Brain and Cognition* **43**, 392–398.
- Stefan, K. M., Fink, G. R., Passingham, R. E., Silbersweig, D., Ceballos-Baumann, A. O., Frith, C. D., and Frackowiak, R. S. (1995). Functional anatomy of the mental representation of upper extremity movements in healthy subjects. *J. Neurophysiol.* **73**, 373–386.
- Steinhauer, K., Alter, K., and Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neurosci.* **2**, 191–196.
- Steinhauer, K., and Friederici, A. D. (2001). Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *J. Psycholinguist. Res.* **30**, 267–295.
- Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain Lang.* **89**, 277–289.
- Stüppich, C., Ochmann, H., and Sartor, K. (2002). Somatotopic mapping of the human primary sensorimotor cortex during motor imagery and motor execution by functional magnetic resonance imaging. *Neurosci. Lett.* **331**, 50–54.
- Stürling, L., and Wales, R. (1996). Does prosody support or direct sentence processing? *Lang. Cogn. Processes* **11**, 193–212.
- Stoffer, T. H. (1985). Representation of phrase structure in the perception of music. *Music Percept.* **3**, 191–220.
- Stromswold, K., Caplan, D., Alpert, N., and Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomography. *Brain Lang.* **52**, 452–473.
- Szymanski, M. D., Perry, D. W., Gage, N. M., Rowley, H. A., Walker, J., Berger, M. S., and Roberts, T. P. L. (2001). Magnetic source imaging of late evoked field responses to vowels: Toward an assessment of hemispheric dominance for language. *J. Neurosurg.* **94**, 445–453.
- Tan, N., Aiello, R., and Bever, T. G. (1981). Harmonic structure as a determinant of melodic organization. *Mem. Cognition* **9**, 533–539.
- Titova, N., and Näätänen, R. (2001). Pre-attentive voice discrimination by the human brain as indexed by the mismatch negativity. *Neurosci. Lett.* **308**, 63–65.
- Thobois, S., Dominey, P. F., Decety, J., Pollak, P., Gregoire, M. C., Le Bars, D., and Broussolle, E. (2000). Motor imagery in normal subjects and in asymmetrical Parkinson's disease: A PET study. *Neurology* **55**, 996–1007.
- Toepel, U., and Alter, K. (2003). How mis-specified accents can distress our brain. Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, Spain, **61**, 9–622.
- Ungerleider, L. G., and Haxby, J. V. (1994). 'What' and 'where' in the human brain. *Curr. Opin. Neurobiol.* **4**, 157–165.
- Vingerhoets, G., de Lange, F. P., Vandemaële, P., Deblaere, K., and Achten, E. (2002). Motor imagery in mental rotation: An fMRI study. *NeuroImage* **17**, 1623–1633.
- Vrba, J., and Robinson, S. E. (2001). Signal processing in magnetoencephalography. *Methods* **25**, 249–271.
- Wang, J., Williamson, S., and Kaufman, L. (1992). Magnetic source images determined by a lead field analysis: The unique minimum-norm least-squares estimation. *IEEE Biomed. Eng.* **39**, 665–675.
- Warren, P., Grabe, E., and Nolan, F. (1995). Prosody, phonology, and parsing in closure ambiguities. *Lang. Cogn. Processes* **10**, 457–486.
- Wernicke, C. (1874). "Der aphasische Symptomenkomplex." Breslau.

- Wheeldon, L., and Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. *J. Mem. Lang.* **34**, 311–334.
- Wilson, S., Pressing, J., Wales, R. J., and Pattinson, P. (1999). Cognitive models of music psychology and the lateralization of musical function within the brain. *Aust. J. Psychol.* **51**, 125–139.
- Winkler, I., Paavilainen, P., and Näätänen, R. (1992). Can echotic memory store 2 traces simultaneously—a study of event-related brain potentials. *Psychophysiology* **29**, 337–349.
- Wimbler, L., Kujala, T., Tittinon, H., Sironen, P., Allen, P., Lehtoleoski, A., Zigler, I., Csepe, V., Jimoniemi, R. J., and Näätänen, R. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology* **36**, 638–642.
- Wolters, C. H., Grasedyck, L., Hackbusch, W. (in press). Efficient computation of lead field bases and influence matrix for the FEM-based EEG and MEG inverse problem. *Inverse Problems*.
- Wolters, C. H., Kuhn, M., Anwander, A., and Reizinger, S. (2002). A parallel algebraic multigrid solver for finite element method based source localization in the human brain. *Comp. Vis. Sci.* **5**, 165–177.
- Wood, C. (1974). Parallel processing of auditory and phonetic information in speech discrimination. *Perception Psychophysics* **17**, 501–508.
- Wraga, M., Thompson, W. L., Alpert, N. M., and Kosslyn, S. M. (2003). Implicit transfer of motor strategies in mental rotation. *Brain and Cognition* **52**, 135–143.
- Zanow, F. (1997). “Realistically Shaped Models of the Head and their Application to EEG and MEG.” PhD thesis, University of Twente, The Netherlands.
- Zatorre, R. J., Meyer, E., Gjedde, A., and Evans, A. C. (1996). PET studies of phonetic processing of speech: Review, replication, and reanalysis. *Cerebral Cortex* **6**(Special Issue), 21–30.