12 *Current approaches to mapping language in electromagnetic space*

Marta Kutas, Kara D. Federmeier, and Martin I. Sereno

'Language , , this great instrument which we have jointly built every word the mystic embodiment of a thousand years of vanished passion, hope, desire, thought.'

Voltairine de Cleyre

12.1 Introduction

The human species is distinguished by its tool-making abilities, and by far the most important and frequently used tool we have created is language. We use language to convey our thoughts and feelings to others via the systematic combination of spoken sounds, manual signs, or written symbols. This ability allows us to bridge the minds of others-sometimes across vast distances of time and space. Language mediates and shapes our social structure, dividing the people of the world according to the kind of sounds or signs they recognize and produce. Language is also used to bring individuals together-to reinforce emotional ties (e.g. wedding vows, business contracts), or to negotiate peace treaties.

Comprehending language requires deriving structure from a stream of auditory or visual inputs at a number of levels. From sensory signals are built phonemes/letters, morphemes, syllables, words, phrases, clauses, sentences, discourses, and, ultimately, concepts. All of these levels are structured-each in its own way-yet extremely flexible as well. Every day we are likely to hear and to produce strings of sounds or signs that we have never before heard or produced. Yet we comprehend these novel streams with ease, perhaps because they are so structured. Having comprehended language input, we can respond by converting concepts into a series of motor commands that will produce systematic changes in our vocal tract and/or our hands.

These transformations require a sophisticated set of sensory receptors and motor effectors. However, whether we comprehend or produce by hand or by mouth, by ear or by eye, the essential processes that allow concepts and feelings to be transferred over individuals, space, and time take place in the human brain. Like language itself, the human brain is structured on a number of levels. On a very gross level, the human brain is made up of two cerebral hemispheres, a thalamus, midbrain structures, a cerebellum, and a hindbrain. These, in turn, can be divided into areas-regions with gross mor- phological or functional. differences. Areas are built from neural ensembles that are

made up of neurons, that are, themselves, just the medium for a complicated set of electrochemical processes.

12.2 Using neurobiological data to understand language processes

Because, in essence, it is the brain from which human language derives, it is to the brain that many turn to discover what language is like, how it is learned and used, and how it might be learned or used more effectively. To do this requires that one look at language in new ways. This chapter describes some of these ways and outlines the improvements that have been made and need to be made in order for the brain-language mappings to yield the information we seek. First, we must begin to understand language as a set of electrochemical processes. Accordingly, we briefly overview the basis of neural communication and outline the strengths and limitations of the various psychophysiological techniques used to measure aspects of electrochemical activity in the brain. These electrochemical processes we measure take place over time and space. Already psychophysiological techniques have allowed us to monitor various aspects of language over time. What has proven harder is the use of these techniques to examine the spatial mapping of language function. We discuss, first, strategies for comparing the spatial distributions of electromagnetic data and for decomposing those distributions into subparts. We then turn to the difficult problem of linking measured distributions to underlying neural generators, outlining the kinds of models used to make this mapping and examining their assumptions.

Turning the measurements we make with our psychophysiological tools into an understanding of the flow of information over space and time and mapping that onto cognition is a difficult and delicate process. Nonetheless, progress is being made and the potential rewards are numerous. By examining the neurobiological roots of language processing, we can ask how the structure and flexibility of the brain mediates the structure and flexibility of language at various levels of both. What brain areas are involved in language processing and what are the more general functions of these areas? What is the extent and temporal order of their involvement in various aspects of language processing? The information that neurobiology has to offer theories of language is very rich and can aid our understanding in a number of different ways. First, neurobiological data can test the psychological reality of the different kinds of language representations posited by linguists and psycholinguists. For example, linguistic theories have suggested that speech is broken down into phonemes, an abstract 'sound' representation that encompasses a number of different possible physical (acoustic and articulatory) patterns. Psycholinguists and linguists also have argued for the existence of a structured lexicon, or 'mental dictionary', of words that mediates associations between their phonological, orthographic, morphological, and semantic aspects. To what extent do we find evidence for these representational levels (e.g. phonological, lexical, semantic, syntactic) and their proposed organizations in the brain?

Neurobiological data also can inform theories about how various representations are used during language production and comprehension. Some theories, for example,

maintain that various language subprocesses are handled by independent, highly specialized 'modules' that are impervious to other types of information. This approach predicts that brain areas processing different types of representations will have little direct influence over one another and will become active in specific sequences (e.g. syntax before semantics). Interactionist accounts, on the other hand, maintain that lower levels of processing/representation are not entirely independent of higher levels but rather interact with them continuously during the processing of a sentence, for instance. Both accounts continue to grapple with questions regarding the domain-generality of language processing. To what extent does the structure of language arise from the functioning of language-specific neurobiological processes and to what extent does it emerge from more general cognitive constraints, such as the amount and availability of attentional and working memory resources?

Neurobiological data can help not only to sort out the nature of language representations and the processes that act on them but also to reveal how language develops, breaks down after trauma or disease, and serves individuals who know more than one. To use neurobiological data to constrain linguistic and psycholinguistic theories, however, requires us to determine what factors the brain is sensitive to and how those factors contribute to language function(s). This is especially difficult, as the concepts and terminology used to theorize about language processing do not readily map onto the terms and concepts used to understand brain functioning. Research using animal models, for example, has shown that factors such as stimulus modality and intensity, frequency (type and token, in the world and in the experimental context), spatial and temporal proximity, similarity (between two stimuli or between a stimulus and a stored representation), and context (physical, experimental, etc.) all influence neurobiological processes. When these factors are manipulated experimentally, one observes changes at various levels; here we give just a few illustrative examples.

Neuronal responses are very context-sensitive; in fact, the tiring pattern of individual neurons to the same stimulus has been observed to change in response to aspects of context that are outside the neuron's 'field of view' (classical receptive field) (Zipser *et al.* 1996). When stimuli are paired together or repeated in rapid succession, the response of neurons may be enhanced for minutes to hours, a phenomenon known as long-term potentiation (LTP) (e.g. original report by Bliss and Lomo 1973); an analogous decrease in neuronal response is observed in some brain areas after inconsistent or infrequent stimulation (LTD, long-term depression) (e.g. Linden 1994). At larger scales, different neurons in a neural ensemble may become active to stimulation depending upon the similarity of the stimulus to something experienced in the distant or recent past. For example, after monkeys had been trained to associate pairs of visual stimuli, researchers noted a significant increase in the number of neurons that would respond to the learned pairs in the inferior temporal cortex of these animals (Sakai and Miyashita 1991). In fact, recent imaging studies suggest that the recruitment of entire brain areas may change with practice (e.g. Raichle *et al.* 1994). These changes can be instantiated with more permanence via the alteration of neuronal connectivity (e.g. Merzenich *et al.* 1988; Black *et al.* 1990). However, it is still relatively uncommon and

often difficult to map between these kinds of data and what is known about language processing.

12.3 Electrochemical basis of neural communication

We can begin to understand how these kinds of factors influence language processing, however, by monitoring neural functioning as individuals perform various language tasks. Various kinds of signals can be monitored, of which the most direct and immediate are electrochemical. Just as humans are distinguished from other species of animals by the complexity of their communication abilities, so neurons are distinguished from other types of cells by their more sophisticated ability to communicate. Channels in the resting neural membrane make it selectively permeable to charge-bearing elements (ions) like potassium, creating a stable electrical gradient between the cell interior and the external fluid. In fact, the electrochemical potential created across the membrane for each type of ion can be thought of as a tiny battery. At rest, a neuron maintains a negative potential (ca. -77 mV) inside relative to outside the cell. Communication between neurons arises from the flow of ions across the neural membrane following changes in the permeability of the membrane. Axonal membranes contain voltage-sensitive sodium channels that open when the membrane potential reaches some threshold, thereby increasing sodium conductance and resulting in a rapid, transient depolarization of the membrane (the inside of the cell becoming more positive). This phenomenon is known as an action potential, or 'spike', and forms the basis of signal transmission in axons.

This disruption of the resting electrical potential travels in wave-like fashion along an axon (as an all-or-none spike) and can be passed on to other neurons via the release of neurotransmitters at synapses. Some neurotransmitters released from the presynaptic terminal open ion channels on the receiving postsynaptic cell that increase the permeability of the postsynaptic membrane to sodium ions, making the inside of the cell more positive. This depolarization of the neuron is called an excitatory postsynaptic potential (EPSP). Similarly, neurotransmitters which increase the permeability of the membrane to potassium and chloride and make the inside of the cell more negative generate inhibitory postsynaptic potentials (IPSPs). Over the past 10 years, a lot of voltage-sensitive channels have been found on dendrites too-perhaps most notable being NMDA channels which need both glutamate and depolarization to open.

The neural communication that underlies human communication thus involves the flow of charged particles across the neural membrane, which generates an electric potential in the conductive media both inside and outside the cell. These transmembrane current flows are the basis both for the electrophysiological recordings in the brain and at the scalp and for the magnetic fields recorded outside of the head for the magnetoencephalogram. These magnetic changes occur in a direction perpendicular to the direction of intracranial current flow, as given by the right-hand rule (if the thumb points in the direction of current flow, the magnetic field points in the direction of the

curled fingers). Viewed from outside the neuron, each patch of membrane acts as a tiny current source or sink, depending on whether the net local current flow is outward or inward, respectively. Both the electric potential and the magnetic field at time t depend on the membrane currents only at that time. This is important for inferences about the timing of electrical and magnetic events at different scalp locations. Moreover, the electric potential field generated by a particular spatial distribution of sources and sinks is the simple linear sum of the individual contribution of each current source and sink in the entire source space.

The development and improvement of psychophysiological tools such as the electroencephalogram (EEG), event-related brain potentials (ERPs), and the magnetoencephalogram (MEG) that are sensitive to these electromagnetic changes is making it possible to begin understanding the flow of information through the brain over time and space. These measures are sensitive to primarily postsynaptic currents, as opposed to the spikes recorded via extracellular single unit recordings' (Ilmoniemi 1993), in brain regions where (i) the average distribution of current sources and sinks within the neurons is distributed in a non-radially symmetric fashion, (ii) the neurons are aligned in some systematic fashion, and (iii) the neurons are activated in local synchrony (see Fig. 12.1).

The neocortex satisfies these constraints. It is organized as a large folded sheet a few millimeters thick wherein about 70 per cent of the cells are pyramidal cells with apical dendrites extending from the soma towards the surface of the cortical sheet. When the proximal parts of these apical dendrites of a cell are activated, current flows preferentially along the length of the dendrite and out of the cell at more distal sites, thereby creating an approximately dipolar source/sink configuration oriented perpendicular to the cortical sheet. Similarly, if the distal parts of a dendrite are activated a dipole field of the opposite orientation is generated. The extracellular currents generated by any single pyramidal neuron are weak, but a cortical region containing hundreds of thousands of such cells activated in synchrony produces a signal strong enough to be detected at the scalp (for more detail, see Kutas and Dale 1997).

EEG, ERP, and MEG techniques measure somewhat different aspects of neural activity, but as a group are among the most direct, non-invasive methods available for the study of neural processing during natural language processing. The EEG measures spontaneous rhythmic electrical activity occurring in multiple frequency bands. Event-related synchronizations in the alpha (S-l 2 Hz) and lower beta (18-30 Hz) bands are taken as electrophysiological correlates of resting or idling cortical areas. Thus, by examining these together with localized, transient attenuation of the EEG activity in the same frequency bands to an event (event-related desynchronizations, or ERD), one can make inferences about the fine structure of neural processing (e.g. Pfurtscheller *et al.* 1994; Krause *et al.* 1996). Another approach examines the average event-related potentials (ERPs) elicited in response to specific events (where 'event' is loosely defined and in some cases refers to preparation for movement or the absence of a stimulus). These are generally measured as a series of positive and negative potential deflections ('components') that can be characterized with respect to their amplitude and latency
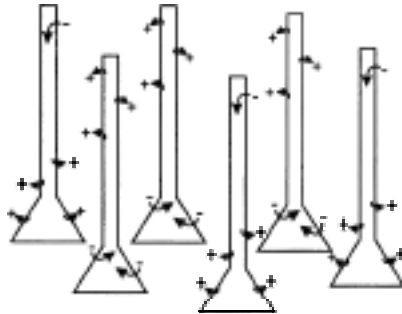
Radially symmetric neurons
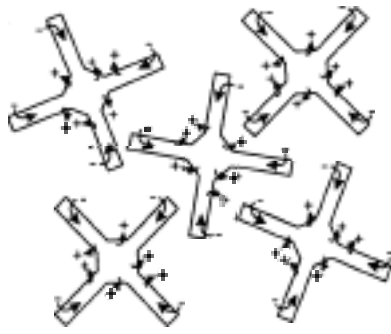
Randomly oriented neurons
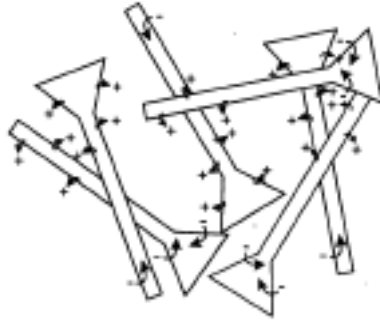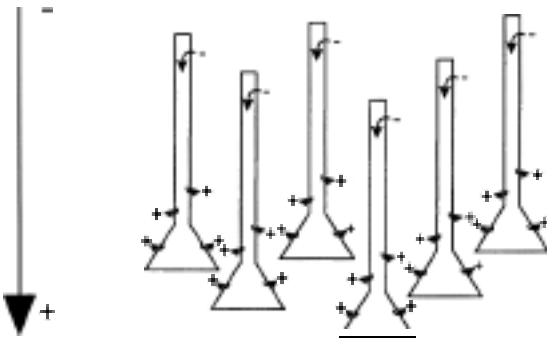
Asynchronously activated neurons

Dipole moment

Fig. 12.1 Examples of closed and open source field configurations. Radially symmetric neurons (top), randomly oriented neurons (second from top), and asynchronously activated neurons (third from top) do not produce externally observable electric or magnetic fields, rather, they produce 'closed fields'. In contrast, neurons that are non-radially symmetric, are spatially aligned, and are activated in synchrony (bottom) produce 'open fields', externally observable electric and/or magnetic fields (From Kutas and Dale (1997) Copyright © 1997 Psychology Press, reprinted by permission.)

across the scalp, although in principle every time point in an ERP waveform can provide valuable information about the ongoing brain activity (e.g. because the surface potential reflects the sum of many differently oriented local sources, a zero potential at a scalp electrode may arise from activity peaks in nearby cortical regions).

12.4 Relationship between EEG and MEG

MEG provides a different, but complementary, picture of the same neural activity measured with EEG/ERPs. As previously mentioned, each local current flow generates a magnetic field orthogonal to the flow (Fig. 12.2). The magnetic fields generated by all active areas add linearly to each other to create the field observed outside the head. The EEG and MEG are affected differently by head shape, dipole location, and dipole orientation. Since so little current flows through the skull, the small magnetic fields generated there can be ignored. Magnetic fields are therefore largely unaffected by inhomogeneities in the skull and intervening tissue (Hämäläinen 1995). With EEG, the skull and the skin must be modelled since the electrical signals we measure pass through them. However, magnetic field strength falls off more rapidly with depth than does the electric potential strength, so MEG is less sensitive to deep sources; furthermore, MEG is insensitive to radial sources (Mosher *et al.* 1993). In practice, this means that the MEG is more sensitive to activity on the banks of sulci and much less sensitive to activity on the crowns of gyri. The EEG is sensitive to both tangential and radial sources, although the electric field due to tangential sources in the fissures may be masked by superficial radial sources (Ilmoniemi 1993).

12.5 Language-related ERP effects

Thus, EEG, ERPs, and MEG allow one to monitor changes in electromagnetic activity coming from. various cortical (and some subcortical) areas. By examining how these changes are correlated with behaviour of interest-recognizing a visual stimulus as a
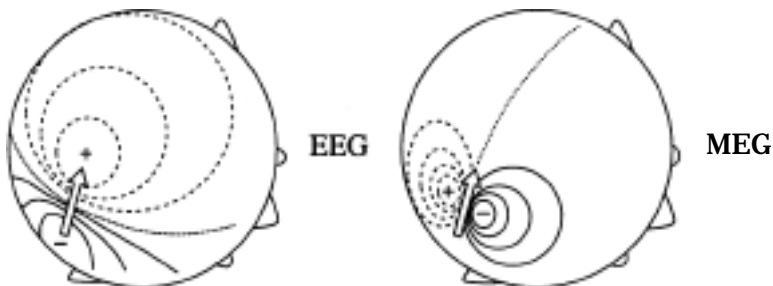

Fig. 12.2   Diagram of hypothetical EEG and MEG field patterns over the scalp produced by a tangential superficial cortical dipole (arrow). Dashed lines indicate positive electrical potential and magnetic flux out of the head while solid lines indicate negative electrical potential and magnetic flux into the head. Magnetic fields are observed perpendicular to electric fields and are less affected by intervening (e.g. skull, scalp) tissue.

word, accessing its meaning, determining its role in the sentence structure-we can begin to map out how cognitive processes are instantiated in the electrochemical medium of the brain. A number of ERP components that change systematically with different aspects of language processing have already been described. Some of these are sensitive to factors (like those already mentioned) that are known to be important for neurobiological processes. Others correlate with variables in language processing that are just beginning to be put in neurobiological terms. Between 200 and 400 ms the ERP to written words over left frontal electrode sites shows a sensitivity to the eliciting word's frequency of occurrence in the language (King and Kutas 1998). The highest correlation ($r = 0.96$) is shown by the latency of a left anterior negativity, referred to as the lexical processing negativity (LPN);[3] this component peaks earlier to the extent that a word is frequent in daily usage. As it overlaps other components in the same time range, it is best seen after digital filtering to remove lower frequency components such as the P2 and N400 (see below). The so-called N280, which has been specifically linked to closed-class or function words (Neville et al. 1992), seems to be an instance of an early LPN due to the disproportionately higher average frequency of this class (which includes articles, prepositions, conjunctions, auxiliaries) relative to members of the open class (including content words such as nouns, verbs, adjectives, and adverbs). While other parts of the ERP vary their amplitude as a function of word frequency, none show a variation in latency. Within the same latency range, the potentials at nearby recording sites show a sensitivity to word length (i.e. number of letters).

The most heavily investigated of the language-related (though not necessarily language-specific) ERP components is a negativity between 250 and 600 ms, peaking around 400 ms post-stimulus onset, with a posterior, right-hemisphere distribution. First described by Kutas and Hillyard (1980) to lexical semantic anomalies, the N400, like the LPN, is a part of the response to every word; in fact without special analytic procedures the two may be difficult to tease apart even though they have different spatial distributions. The N400 seems to be the default response to words whether they occur in auditory speech, in sign language, or in written text (as long as they are orthographically legal). The amplitude of the N400 to content words in lists is smaller for those that are abstract than concrete and decreases with word frequency, repetition, and orthographic, morphological, phonological, and semantic priming; in sentence contexts, its amplitude decreases as the predictability of the word in the context increases ('cloze probability', usually due to increasing contextual constraint) (see Kutas and Van Petten 1994, for more detail, Fig. 12.3 gives an example). The reduction in N400 amplitude across the course of a sentence has been interpreted as reflecting a reduction in the difficulty of sentential integration due to the build-up of semantic constraints (Van Petten and Kutas 1990).

Note that the presumed equivalence of N400s across sensory modalities is based on functional rather than spatial similarity. For instance, while the amplitude of the N400 to both a written and spoken word is reduced by a preceding semantic associate, the visual N400 is later and more prominent over the right posterior areas than the auditory N400. In fact, even within a sensory modality, the apparent distribution of an

They wanted to make the hotel look like a tropical resort.
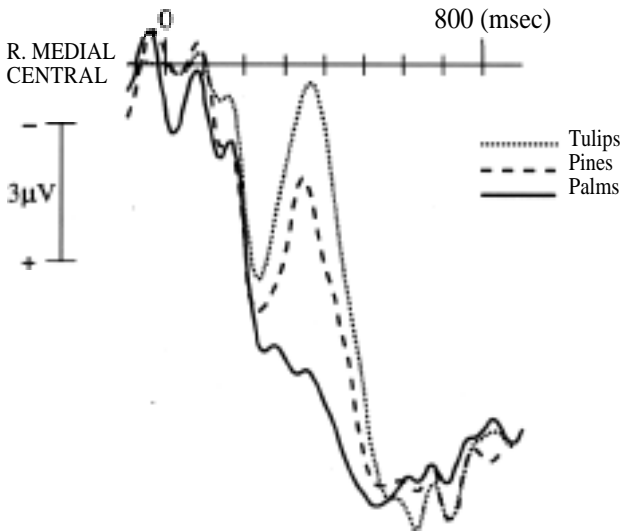So along the driveway they planted rows of...

Fig. 12.3 ERPs to visually-presented sentence-final words, shown at a right, medial electrode site. The response to words expected in the context (solid line) is characterized by a sustained positivity. In contrast, ERPs to words unexpected in the context (dashed and dotted lines) are characterized by a negative-going potential peaking approximately 400 ms post-stimulus onset (N400). In addition to its sensitivity to sentential context, the N400 effect reflects long-term memory structure; responses to unexpected words from the same category as the expected word (dashed line) show a reduced N400 response relative to equally unexpected words from a different semantic category (dotted line). (Data from Federmeier and Kutas, submitted.)

N400 effect (difference between responses to congruent and incongruent words, for example) may change with factors such as speed of presentation; faster presentation rates are associated with more frontal negativity in the N400 region than slower rates (e.g. Kutas 1987). Such changes in the surface distribution of the potential imply a change in the orientation or location of the active neural sources, which implies that different parts of the cortex are generating the signal.

Electrophysiological researchers have identified two ERP components that seem to be sensitive to syntactic manipulations (for a more complete review, see Hagoort *et al.* Chapter 9). One is a negativity, called the left anterior negativity or LAN, whose latency range is similar to that of the N400 but which has a more anterior distribution and left hemisphere bias (Neville *et al.* 199 1; Osterhout and Holcomb 1992; Rösler *et al.* 1993). Münte *et al.* (1993) used distributional as well as functional differences to argue that the LAN effects are distinct from the N400 component, are produced by different underlying generators, and index syntactic aspects of processing. Kluender and Kutas (1993*a*; 1993*b*), however, have suggested that the LAN elicited by certain syntactic violations actually indexes some aspect of working memory usage. Even in grammatically correct sentences, Kluender and Kutas find a LAN effect that is

associated with entering a filler in working memory, storing it, and subsequently retrieving it to assign fillers to gaps. King and Kutas (1995) likewise observed a LAN effect in ERPs elicited by verbs that tax working memory; specifically, they recorded a LAN to a main-clause verb immediately following the gap in object-relative clauses in comparison to the corresponding verb in subject-relative clauses, which in turn had a larger LAN than verbs in unembedded sentences. As this class of negativities does show some variation in distribution, timing, and especially in its degree of lateraliza-tion, it may not reflect a unitary process.

   The syntactic effect that has received greater attention is a slow, positive shift observed in response to violations of a range of syntactic phenomena, including agreement, phrase structure, subcategorization, and subjacency (Neville *et al.* 1991; Osterhout and Holcomb 1992; Hagoort *et al.* 1993; Münte *et al.* 1997; Coulson *et al.* 1998). This positivity has variously been labelled the P600, or the syntactic positive shift (SPS). The nature of the P600 component has not been wholly consistent across studies. For example, Osterhout and Holcomb (1992) report a P600 in response to both phrase structure and subcategorization violations, but they vary in scalp distribution. Moreover, for the same sort of phrase structure violations that yield a positivity with right anterior distribution, Neville *et al.* (199 1) report a laterally symmetric positivity largest over occipital regions, Likewise, for the violation of subcategorization con-straints to which Osterhout and Holcomb (1992) report a positivity with symmetric posterior distribution, Hagoort *et al.* (1993) find no effect. This variation may reflect the existence of more than one late positivity with different scalp distributions in this time interval, or the overlap of other components (such as the LAN or N400) which may alter the apparent distribution at the scalp. The P600 is typically described as beginning around 500 ms and having its midpoint around 600 ms, with a somewhat posterior maximum. This component, or set of components, seems to be sensitive to grammatical violations, both locally (e.g. agreement) and more globally (e.g. phrase structure). The fact that it is also sensitive to the probability of the violation in the experimental context suggests that it may be related to the family of positivities that includes the P300 (see Coulson *et al.* 1998).[3]

   In addition to analysing these transient ERP effects elicited by linguistic violations or regions of lexical, semantic, or syntactic ambiguity, researchers have begun to monitor potentials time-locked to entire clauses or sentences, or larger parts than individual words therein. Moreover, many of these investigations have focused on wholly congruent sentences (i.e. without linguistic violations of any sort) that vary only in their structure (e.g. number of embeddings). These cross-clause potentials tend to be of a lower frequency than the transient effects and are often best seen after a low-pass digital filter is applied to the raw ERP data. Such filtering reveals that there are elec-trophysiological measures that emerge across sentences that are more than the responses to the individual words lined up one after another (Kutas and King 1996).

   These slow potentials show systematic variation in time across the extent of the clause and in space across the scalp in both the anterior-posterior and lateral dimensions (for review, see Kutas and King 1996). For example, a sustained negativity

over occipital regions is insensitive to word class but is specific to the processing of visual (as opposed to auditory) features. Anterior temporal sites, by contrast, do show a sensitivity to lexical class. A phasic positivity primarily over the left hemisphere is observed coincident with a verb's occurrence in a sentence, both in word by word reading and natural speech. It has been hypothesized that this positivity reflects some aspect of thematic-role assignment based on information contained in the lexical representation of the verb. Associated with clause endings during reading and listening, these slow potentials are characterized by a clause ending negativity (CEN), which is also somewhat better defined over the left than right temporal and central sites; these have hypothetically been linked to working memory operations at clause boundaries, so-called wrap-up processes. Perhaps, the most striking effect in these cross-clausal data is an extremely slow (< 0.2 Hz), cumulative positivity over (especially left) frontal recording sites. It has tentatively been linked to some executive function(s) of working memory such as the integration of items in working memory with information from long-term memory. Thus it might reflect the use of long-term memory to build a mental model, schema, or frame (message level representation) of the incoming sentence in working memory.

We examined this hypothesis further by contrasting two sentence types that vary in working memory demands by virtue of the differences in conceptual knowledge activated by their initial words (*before* vs. *after*) (Münte *et al.* 1998). Our real world experiences suggest to us that time unfolds sequentially, with current events sometimes causing future events. Our linguistic knowledge tells us that temporal conjunctions often draw attention to the sequence of events in a discourse. Moreover, whereas *after* signals that events will be expressed in their actual order of occurrence, *before* signals that events will be expressed in reverse order. We believe that there should be processing consequences of this. In a sentence beginning with *before,* the first clause cannot be fully integrated in a message level representation until after the second clause. In a sentence beginning with *after* the first clause can be integrated upon its completion. The former clearly places more demands on working memory. As can be seen in Fig. 12.4, the ERPs to the two sentence types diverge soon after the initial words; the waveform to the less-demanding or easier to integrate *after* sentences goes positive whereas that to the more taxing *before* sentences stays negative (see also King and Kutas 1995 for a similar effect comparing subject- versus object-relative sentences).

The different spatial distributions of these various effects also point to the distributed nature of aspects of sentence processing. By continuously recording across clauses and applying low-pass digital filtering, it is possible to monitor some of the overlapping but different processes that take place in multiple brain regions at the same time, albeit with different time courses. Also of note is that many of these slow potential effects associated with reading of, or listening to, sentential clauses differ reliably as a function of the comprehension skill of the reader or listener. Good comprehenders, for instance, show larger occipital negativities and smaller ultra-slow frontal positivities than poorer comprehenders. In summary, there are a number of fast and slow ERP responses to word- and clause-level effects that vary systematically in their timing and
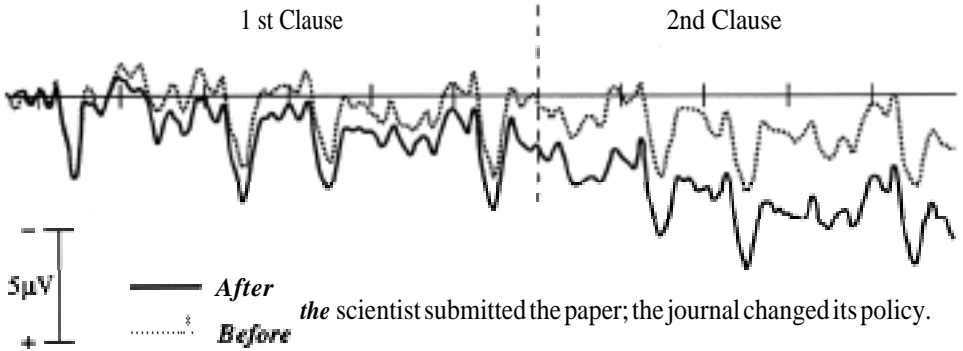
1 st Clause   2nd Clause

5μV

——— *After*

.... *Before*

the scientist submitted the paper; the journal changed its policy.

Fig. 12.4   Over-sentence ERPs to two sentence types recorded at a left frontal electrode   as volunteers (*n* = 8) with high verbal working memory span scores read sentences one word at a   time for comprehension. The clause boundary is marked. Single word responses (repeated,   higher-frequency activity) can be seen riding on top of a slow, sustained difference in which responses to sentences beginning   with 'after' (solid line) are progressively more positive than (otherwise identical sentences) beginning with 'before' (dotted line) (From Münte, *eta/* (1998). Copyright © Macmillan Magazines, reprinted with permission)

spatial distribution in ways that can be used to answer psycholinguistic questions at the level of word recognition as well as discourse.

12.6 Issues in ERP signal analysis

12.6.1 ERPs and the time course of processes

Thus, language components have been differentiated from one another and tied to cognitive processes via both their temporal and spatial properties. Of the two, the measurement of time using EEG/ERPs/MEG is more straightforward. Because these are direct, real-time measures of neural electrical activity, their temporal precision is quite high, with an upper limit at the sub-millisecond level. This precision is given by the fact that both the electric potential and the magnetic field at time t depend on the membrane current at time t only; in other words, the propagation of the potential and magnetic field is essentially instantaneous. Thus, if t is the earliest time at which ERPs from the two conditions differ significantly, it can be concluded that the brain activity differs between the two conditions at that time. The onset of the latency of the ERP difference between two conditions can thus be taken as an upper limit on the time by which the brain must have processed the stimuli sufficiently to distinguish them. Note that the converse does not hold; there are many reasons why one might fail to detect a difference between two conditions.

The order of processes for some language function and the duration of any given process can be critically important given the long-range dependencies that characterize language comprehension. In fact, many of the current debates about language processing revolve around issues of timing such as when some information becomes available or when it is used. Timing is of the essence when considering, for example, the nature of lexical access (is it automatic or controlled?), the extent to which

syntax is autonomous, and the influence of context. Many language processes are extremely fast-individuals are able to produce words at a rate of two to three words per second (Levelt 1989)—and to follow these processes it is necessary to have a real-time measure that has a temporal resolution on the order of milliseconds. Other language processes, however, are quite slow because they must span a long discourse, a lengthy passage, or a sentence, which, given the recursive nature of language structures, could also be quite long. Indeed, the practical beauty of the ERP methodology is its applicability to language processes of any duration. Electrophysiological techniques thus provide the only dependent variables that can span both the crucial milliseconds determining what phoneme has been uttered and the seconds or minutes that may be needed to determine who did what to whom in a wh-question or a sentence with one or more embeddings (e.g. *The cat the rat the bat bit saw lunged and all hell broke loose temporarily*).

12.6.2 ERPs and function@ independence

Since the electromagnetic activity associated with different generators adds linearly, ERP and MEG methods can also be used to determine the extent of functional independence between different brain mechanisms. If two conditions/processes are completely independent-that is do not share neural resources-then the activity generated when the two processes occur and overlap in time will be the sum of the activity generated by each process alone. If a given source is involved in both conditions, there may be non-linear interactions; for example, one condition may already drive a source to its maximum output. Nevertheless, we can count on the output of any given source adding linearly to the signal we observe from other sources.

Thus, the ERP can be used to examine when and how the brain subdivides complex cognitive tasks. For example, Besson and her collaborators (manuscript) performed a series of experiments to determine if musical lyrics and musical tones are processed independently or form a single percept. While semantically incongruous words within sentences elicit an N400 component (e.g. Kutas and Hillyard 1980), harmonically incongruous notes are associated with a positivity (the late positive complex, LPC) in a similar time window (Besson and Faita 1995). To examine whether these differences reflect independent processing of musical tones and musical lyrics, Besson *et al.* recorded ERPs as professional musicians listened to four types of operatic excerpts: (i) semantically congruous and sung in key, (ii) semantically congruous but sung out of key, (iii) semantically incongruous but sung in key, and (iv) semantically incongruous and sung out of key. As noted in previous studies, semantically incongruous items elicited an N400 while out-of-key notes elicited an LPC. Results in the double incongruity condition were consistent with an additive model of these two components, arguing for a complete independence of semantic and harmonic processing.

12.6.3 The value of multiple recording sites

Determining that the brain is sensitive to a particular factor and following this sensitivity over time could, in principle, be done by recording from only a single electrode.

Since electrical and magnetic fields are generally very small (electrical activity at the scalp is on the order of tens of microvolts and the magnetic fields recorded are nine orders of magnitude weaker than the earth's magnetic field), at times it proves useful to record from multiple sites to verify the reality of the measured signal change (though, in principle, an effect can be limited to one location and still be real; the LPN, for instance, is quite focal). Replication and averaging, however, can be as effective as more recording sites for this purpose. On the other hand, if the data suggest that some variable has no effect then more channels may be needed to verify this conclusion, since it is always possible that recordings were made at the wrong sites and/or that the effect has been missed for some other reason.

Traditionally, EEG recordings have been based on the 10–20 system of electrode placement (Jasper 1958; also see Böcker *et al.* 1994). It has the advantage of being standardized (and has been routinely used in the clinic). However, electrode placements are relatively sparse, with interelectrode distances on the order of 7cm (for reference, the distance between two adjacent cortical gyri is on the order of 1 cm). The Nyquist distance[4] for cortical generators of electrical signals dictates interelectrode distances on the order of 3 cm or less (closer to the spacing of gyri) (Spitzer *et al.* 1989). Yvert *et al.* (1996) found significant improvements in dipole localization accuracy in simulated data by increasing the number of electrodes from 19 to 63. Equally as important as the high density of recording sites is that they be systematically placed on the head. The best would be an equidistant placement of the electrodes covering the entire surface of the head (including the bottom portions). A subtesselated icoscahedral configuration provides excellent uniform coverage of the top of head. It is also critical, however, that electrode impedances at the scalp be low (below 3 kOhms) and identical to one another.[5] For some uses (related to increasing spatial resolution), it is also important to know the exact locations of the electrodes. The average error in standard electrode placement is estimated to be on the order of 1-2.5 cm (Kavanagh *et al.* 1978). Methods for direct localization of electrode co-ordinates, such as with three-dimensional digitizers, are available (e.g. Echallier *et al.* 1992). Under conditions of high spatial sampling and appropriate recording conditions, the absence of an effect may form the basis of a working hypothesis that the generators of the scalp potentials are probably not differentially sensitive to the variable in question.

Determining that the neurons creating the activity measured by the dependent measure are not sensitive to some variable, however, is not the same as determining that the brain is insensitive to that variable. Not all brain events can be seen at the scalp surface. For example, active neurons may be radially symmetric, or may be arranged in a closed field, or may become active too asynchronously to signal their involvement at the scalp (see Fig. 12.1). Since not all brain activity can be seen at the scalp's surface, it is always more difficult to interpret the absence than the presence of an effect. Nonetheless, given what is known about the primary generators of the ERPs, the lack of an effect in clean, replicable data could lead to the hypothesis that at least the pyramidal cells in much of the neocortex are probably not sensitive to the manipulated factor. The fact that some brain activity cannot ever be seen at the surface adds to the

inherent limitations in using the ERP to localize, given that the same spatial pattern of potentials at the scalp can be consistent with an infinite number of source configurations in the brain. We will discuss this so-called inverse problem and ways to deal with it in more detail later.

In most cases, when we do find an effect we are not merely interested in its existence or even just its timing. Rather, we would like to be able to identify it-to link it to and differentiate it from other known ERP effects. For this purpose, data from one recording site definitely will not suffice. An effect can vary in latency as a function of stimulus, response, and cognitive variables and still be generated by the same source configuration. Thus, many components/effects can only be definitively identified by virtue of their scalp distribution, although this is not easily accomplished. The N400, LPN, and LAN all have known characteristic distributions, and it is as much their different distributions as their different timings/temporal courses and sensitivities to experimental variables that identities a negativity as an N400 and not an LPN or an LAN, for example. If we could assume that the mind was a serial stage processor, it would be relatively easy to characterize the spatial distribution of the associated ERP effect by measuring the amplitude of the potential at all recording locations at a given time point (peak amplitude) or within some range (area or mean amplitude). In reality, however, there is enough empirical evidence for the existence of cascade-like and parallel cognitive processes to render this simple serial assumption invalid. Mental processes and the associated ERP effects overlap in time and therefore in space, making it analytically quite difficult to define the spatial distribution of an effect unequivocally.

12.6.4 Comparing ERP scalp distributions

Given the importance of distributional analyses for many kinds of questions, it is a major analytic shortcoming in the field that there are no commonly agreed upon, objective techniques for comparing scalp distributions. These comparisons are typically made using electrode as a variable in a repeated factors ANOVA or by breaking down electrode locations into laterality (left vs. right, sometimes also including medial vs. lateral vs. midline) and anteriority (anterior-posterior with several levels) and using those as variables. McCarthy and Wood (1985) addressed the potential problems raised by such comparisons when the amplitudes of the effect for which distributional differences are being assessed are unequal. They pointed out that amplitude differences can be misinterpreted as differences in distribution and suggested normalizing the measures and performing the ANOVAs on normalized data. If a reliable interaction between some experimental manipulation and an electrode factor remains after such normalization, the inference that there is a distributional difference is justified.

The better the characterization of the spatial distribution, the more accurate our ability to differentiate two components at the level of scalp distributions or to determine that different neural generators are involved in different experimental conditions. A number of factors have led to the poor spatial resolution of conventional EEG/ERP methods at the scalp and in the brain. These include limited spatial sampling, reference electrode contamination, and failure to use inverse methods including information

about the volume conductor (see Nunez *et al.* 1993). All of these shortcomings have been addressed in the recent past, however, and the current spatial resolution of the EEG/ERP is much better than it is typically characterized as being.

As already mentioned, the use of more electrodes, under the right conditions, can lead to a much improved characterization of the spatial distribution of the signal (also see later section on the forward solution). However, the utility of increased spatial sampling is ultimately limited by the distortion of the neuronal potentials as they are conducted through the highly resistive skull. The appreciably larger resistivity of the skull than the brain leads to current flow parallel to the skull for a distance, effectively smearing the image of a cortical source. The practical consequence of this is that the use of more than 256 electrodes (approximately intergyrus spacing) is unlikely to provide much additional information. The smearing of the potential distribution by the skull differentiates the EEG from the MEG. Because of the nature of the MEG forward solution (cancellation of magnetic fields generated by return currents in a spherical conductor), MEG field maps typically contain somewhat higher spatial frequencies than corresponding EEG potential maps; it has been suggested that this may be an advantage for recovering more information about multiple or distributed neural electromagnetic sources.

12.6.5 The Laplacian or current source density

One common deblurring or spatial enhancement technique is the Laplacian operator, which is the second spatial derivative of the potential field at each electrode.[6] The Laplacian is also called the current source density (CSD) and is proportional to the current entering and exiting the scalp at each electrode site (if the cortex was actually on the surface of the head and there were no deep sources, a high resolution CSD map would exactly describe the pattern of current flow into and out of the cortex) (e.g. Nunez 198 1; Perrin *et al.* 1987). The Laplacian measures the curvature of the potential field, which is often not easy to see by inspecting the potential map. For example, a generally positive region in the potential map may signify either a source, a sink, or no current flow depending on its curvature. In a potential map, a small source or sink can easily be overwhelmed by a more broadly distributed potential on which it rides. By contrast, these functionally significant changes in the field are revealed in a CSD map (e.g. Fig. 12.5). Laplacian estimators of scalp current density are reference independent and converge on the true spatial pattern of scalp current sources and sinks as the density of electrodes increases.

However, because the Laplacian is a derivative operator (i.e. based only on nearest neighbour differences), it is sensitive to noise. Typically, the potential field is first interpolated with a smooth function before the derivative is taken. One typical function for this purpose is a spherical spline, which allows interpolation of data from irregularly spaced electrodes (Perrin *et al.* 1989; also see Nunez and Westdorp 1994). The CSD is relatively insensitive to signals that are common to the local group of electrodes used to compute it; thus it is more sensitive to cortical potentials of higher
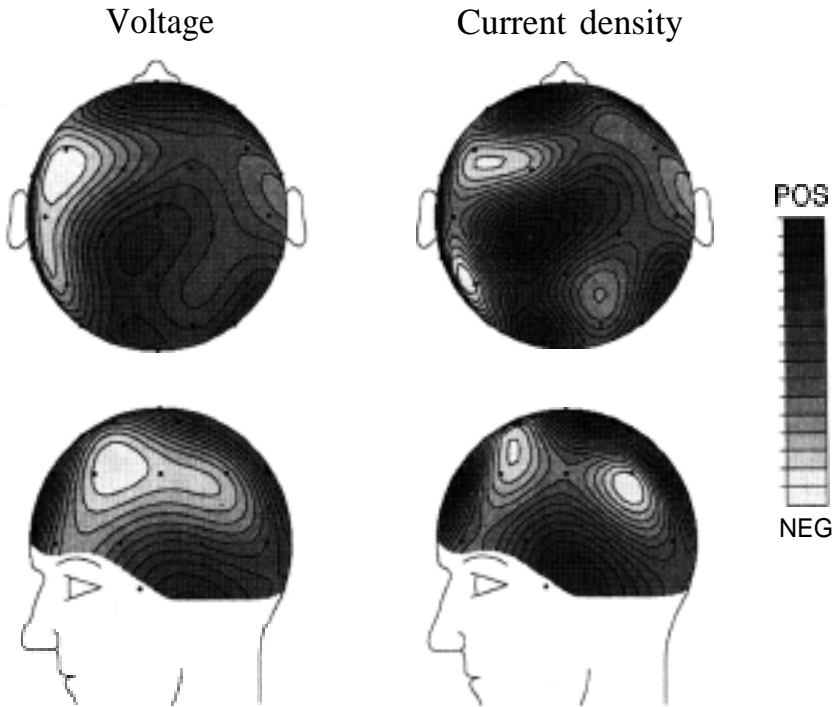
Voltage

Current density

Fig. 12.5 Comparison of Interpolated voltage and current source density maps Shown are top and left side views of electrical activity recorded as a group *(n=* **24)** of individuals read sentences one word at a time for comprehension Plotted is the mean amplitude of difference in activity recorded across two sentence types that differ only in their initial word [e.g 'Before/After the scientist submitted the paper, the journal changed its policy'] Note that the *Before* sentence is associated with greater negativity over left frontal sites (From Munte et al unpublished data)

spatial frequency than to signals conducted from sources (subcortical) further away (e.g. Gevins 1996).

Although accurate distributions are important, at times inappropriate inferences are drawn from these distributional analyses. For example, when it is said that the visual N400 has a posterior, right hemisphere dominance, this means that it is generally larger over posterior rather than anterior electrode sites and generally larger over the right than the left hemisphere (at least when certain references are used, such as linked or average mastoids or non-cephalic reference). While this description has some heuristic value, it can be misleading—first because people may tend to believe that the N400 effect at the scalp represents the activity of some single source (neural generator) within the brain and second because they might assume that the fact that an effect has the largest amplitude over some scalp location reveals the source of its underlying generator in some transparent way. However, the scalp N400, for example, is likely to reflect activity from several neural generators (see intracranial discussion below). Additionally, very deep, midline sources can have quite broad distributions at the

scalp, from which little information can be gleaned about the actual location of the generator(s). Moreover where on the scalp a potential is maximal depends very much on the orientation of the generator. For example, it is well known that generators on the medial surface of a hemisphere can be 'paradoxically' lateralized-that is, be larger over the contralateral hemisphere (e.g. Regan 1989).

Neural activity in a particular location in the brain produces a potential field with a characteristic scalp distribution. Simultaneous neural activity in two different loca-tions would result in a scalp distribution that would be the sum of the characteristic scalp distribution of each alone, and so on. Thus, the distribution of activity at the scalp at any given moment is not a pure measure of any process unless we can be certain that only one generator is active. It is often assumed that a peak reflects the same process at all electrodes, and that differences in peak latency across the scalp reflect propagation of the process from one side of the head to the other. In fact, changes in scalp distribution over time necessarily imply that several generators with different time courses are involved. Because of the nature of electrical field propa-gation, neural activity that overlaps in time will also overlap in the spatial distribu-tion of the effect and be quite difficult (if not impossible) to identify without quantitative modelling.

12.6.6 Extracting signals

Since different brain generators produce widespread, overlapping potential distribu-tions, it is essential to have a method for extracting the signal produced by each gen-erator. If the scalp distribution of each generator were known, then the contribution of each could be determined using a weighted sum of the potentials recorded at each electrode, with weights given by a least squares solution. In other words, at any given moment the signal recorded in, for example, an ERP experiment can be thought of as composed of the sum of a number of independent electrical processes. For instance, an individual may be simultaneously processing a semantic incongruity, processing a harmonic incongruity, shifting position in her chair (generating muscle activity), and blinking her eyes. These signals may be independent, but because they overlap in space and time it can be quite difficult to determine which aspect of the measured waveform is due to which of the possible sources.

In a recent paper, Bell and Sejnowski (1995) describe a new unsupervised learning algorithm that maximizes information transfer from input to output through a net-work consisting of non-linear units. Because the units that implement the input-output transfer function are non-linear, the system is able to pick up on higher-order moments of the input distributions (that is, non-linearities) and reduce redundancy between units in the output representation. The consequence is that the network separates statistically independent components in the inputs. This technique, used on the hypothetical data described in the previous experiment, would return the independent components of the original waveform-dividing them into a waveform due to the processing of semantic incongruity alone, that due to the processing of the harmonic

incongruity alone, that due to the muscle activity alone, and that due to the blink. These independent components can be linearly recombined to reconstruct the original input, or they can be manipulated independently-used, for example, as templates for removing muscle and eye blink artefact from the original waveform.

This procedure is related to PCA-principal components analysis-as well as other separation methods such as generalized eigen vector analysis (Dale 1994) that use spatial filters to reduce the dimensionality of ERP data to a smaller set of uncorrelated or independent components. At a minimum, these procedures make it easier to separate the signal of interest from artefacts (such as the electrical signals caused by blinks or heart and striate muscle activity). But such procedures also may be quite effective at separating signals with distinct spatial distributions and therefore useful for decomposing waveforms in terms of informative experimental manipulations (for spatial PCA see, e.g. Skrandies and Lehmann 1982). Note, however, that blind application of such methods without due concern for experimental manipulations and/or the associated neurophysiology can be misleading.

Another new technique involving iterative deconvolution, first introduced by Woldorff (1993), can also be used to tease apart overlapping responses (for example, that to a stimulus versus that to a response). Such overlap is commonplace, occurring whenever the interval between adjacent events is smaller than the duration of the event-related response. Overlap of this sort can lead to significant distortion of the average event-related ERPs. Woldorff's procedure works only with moderate amounts of overlap and does not provide an exact solution. Dale *et al.* (in press) have provided a general exact solution to the overlap problem called DOC-a direct method for overlap correction (also see Hansen 1983).

12.7 Localizing ERP generators

Although a better specification of an ERP effect's distribution can help resolve some kinds of issues and answer some kinds of questions, for the purposes of functional brain mapping we are often interested in determining where in the brain an effect is taking place. ERP data can contribute to this process in several ways. First, while it clear that ERPs can provide useful information about the temporal course of language-related processing, it may be much less obvious that timing information can also be very important for the purposes of localization. It is generally assumed that localization of active language areas requires a brain measure with high spatial resolution such as PET or fMRI. But spatially contiguous activity may be hard to tease apart except by temporal course differences. Moreover, it may be that some functional activity (e.g. phase-locked oscillatory activity) can only be defined by its temporal characteristics and thus requires methods such as ERPs/MEG (or fMRI) for its resolution. Since electromagnetic measurements are also sensitive to the orientation of current flow, they can aid in distinguishing nearby, but functionally distinct, regions of cortical activation.

12.7.1 *intracranial recordings

In some cases, the localization of ERP components recorded at the scalp has been aided by intracranial recording in humans. In these experiments, recordings are made via subdural electrode grids or depth probes in patients undergoing evaluation for medically intractable epilepsy (e.g. Lesser *et al.* 1994; Fried *et al.* 198 1). Relatively few intracranial experiments have been done looking at language sensitive components; recently, however, a handful of studies examining word repetition, semantic priming, and sentence context effects have been reported. Building on earlier work by Smith *et al.* (1986), McCarthy *et al.* (1995) examined intracranial responses to congruent and anomalous words in sentence contexts. Anomalous words were associated with a bilateral negative field potential in the anterior medial-temporal lobe with peak latency near 400 ms (Fig. 12.6). Electrodes near the collateral sulcus recorded positive field potentials at the same latency. Because a potential arising from a particular source will be positive when measured from one direction and negative when measured from the other, this kind of 'polarity reversal' can be used to narrow down the spatial location of a source; in this case, the authors suggest that the voltage pattern is consistent with a source in or near parahippocampal and anterior fusiform cortices.
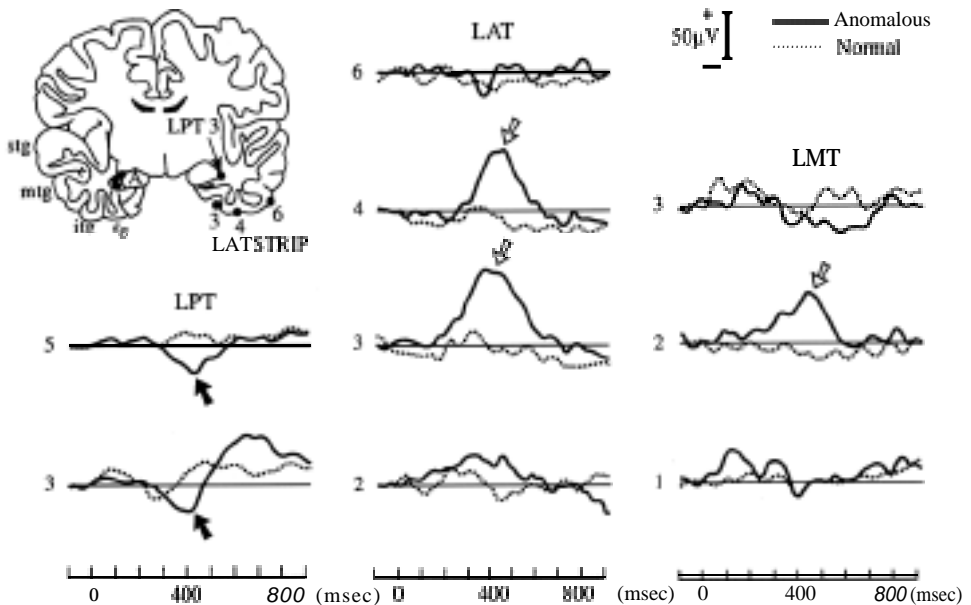

Fig. 12.6 ERPs elicited by anomalous (solid line) and normal sentence-ending words recorded intracranially. The black arrows point to the anterior medial temporal lobe (AMTL) N400, which was recorded from the left posterior temporal (LPT) probe. LPT3, shown on the MR tracing, was located on the lateral border of the amygdala and is indicated by the labeled arrow. The white arrows point to AMTL P400s which were recorded from left anterior temporal (LAT) and left medial-temporal (LMT) subdural strips. Polarity inversions like these can be used to determine approximate locations of the neural generators of the measured response. (From McCarthy *et a/.* 1995. Copyright © 1995 Society for Neuroscience, reprinted by permission.)

The component observed in the anterior medial-temporal lobe to anomalous words in sentences shares several important properties with the scalp-recorded N400 component already described. Isolated words generate these potentials, but only if they are orthographically legal (Nobre and McCarthy 1995). An earlier intracranial component called the N200 recorded in the posterior fusiform gyrus is sensitive to letter strings regardless of their legality (Allison *et al. 1994;* Nobre *et al.* 1994). The later potential, like the N400, is larger to content words than to function words and is diminished by semantic priming (Nobre and McCarthy 1995). Like the N400, this component also seems to be associated with memory, though the potential generated during, for example, continuous recognition tasks may be more widely distributed among medial temporal structures including hippo-campus and amygdala (Guillem *et al.* 1996). Some have argued for a functional dissociation across structures for the roles played in memory-related processing by potentials in this time window, with left medial-temporal N400s predicting immediate recall performance and left anterior medial-temporal N400s predicting delayed verbal recall (Elger *et al.* 1997).

Even with intracranial recordings, however, precise localization can be difficult and the mapping between depth potentials and surface potentials is not straightforward because of the infoldings of the cortex. Volume conduction is also a problem for localization in depth, and, in the absence of a dense three-dimensional grid of recording sites, it can be difficult to know from where the source of the electrical activity recorded at a particular electrode emanates. That is, depth recordings are also subject to the inverse problem, discussed below. Additionally, electrical activity recorded at the scalp may arise from the summation of many, divergent brain sources. For example, while the anterior medial-temporal component behaves much like the scalp-recorded N400, the authors note that anomalous endings are frequently, though not consistently, associated with activity that seems to be generated in the hippo-campus proper (McCarthy *et al. 1995).* Studies of memory have noted activity in the N400 latency range in the frontal, parietal, and occipital lobes, as well as the temporal lobe (Guillem *et al.* 1995), and it is unknown whether these findings suggest a dis-tributed source for the N400 or merely reflect volume conduction of more localized electrical sources. Thus, while intracranial studies provide valuable information, they do not, in and of themselves, allow for exact localization.

While ERP/EEG data can contribute to localization via temporal information and while intracranial techniques can bring us closer to the source of scalp-recorded-potentials, we would ultimately like to use our EEG/MEG measurements to specify the current distribution throughout the brain volume. In other words, we would like to know the number, location, spatial configuration, strength, and time course of neu-ronal currents that give rise to the potential distribution or magnetic field we record at the scalp (or in some cases in the depths). In short, we are faced with the problem of calculating current distributions inside the brain given magnetic field and/or scalp potential measurements on the outside; this is known as the electromagnetic inverse problem.

12.7.2 The forward solution

Before discussing why the inverse problem is inherently insoluble, let us quickly examine the electromagnetic forward problem, whose solution is limited only by the suitability of the source and head models chosen. Calculating the electrical potential and/or magnetic fields outside the head given a particular distribution of current inside is known as the forward problem. Its solution depends on a source model of the properties of the current sources (including location, orientation, and amplitude) and a volume or head model of the electromagnetic properties of the brain, skull, and other tissues as electrically conductive media (Fig. 12.7 gives an example). A variety of source (single point current dipoles, dipole sheets, and realistic current distributions) and head (single homogeneous sphere, multicompartment spherical shells, realistic) models have been tried.

12.7.2.1 Head models

A commonly used head model consists of three concentric spheres representing the boundaries between the air/skin, skin/skull, and skull/brain (sometimes the brain and cerebrospinal fluid are distinguished using another sphere). Such shell models do not provide variations in the conductivity of the scalp, skull, and brain and assume a single value for electrical conductivity. Thus, spherical shell volume conduction models are only a first approximation for electrical measurements, as the shape and thickness of the low-conductivity skull and, actually more importantly, the skin (on which the electrodes are placed) can have significant effects on the surface potential data (Cuffin *et al.* 1991). Finite element models take into account various geometric factors such as the shape of the skull, variation in its thickness, presence of skull openings (foramina and sutures), as well as the location, size, and shape of the ventricles. They use a large number of finite volume elements to create a head model which is anatomically accurate to any arbitrary level of precision.

Using high resolution MRIs, some have demonstrated three-dimensional finite element methods that can model the entire head volume at several millimetre resolution, accounting for fluid-filled spaces, local inhomogeneities, and anisotropies arising from white matter fibre tracts (George *et al.* 1995). While finite element deblurring (FED) methods can clearly increase the available spatial detail, the improvements do not come without a price. They require considerably more knowledge about the electrical (resistivities) and geometric (tissue boundaries) properties of small parts of the head. This information is not available from MRI without additional assumptions about the relation between MRI contrast and conductivity. Direct methods for measuring local conductivities (electroimpedance tomography) have extremely course resolution, for now. Finally, FED becomes computationally intractable with small elements.

A compromise between spherical analytical solutions and finite element methods is the boundary element method. Like the spherical shell model, it assumes that the
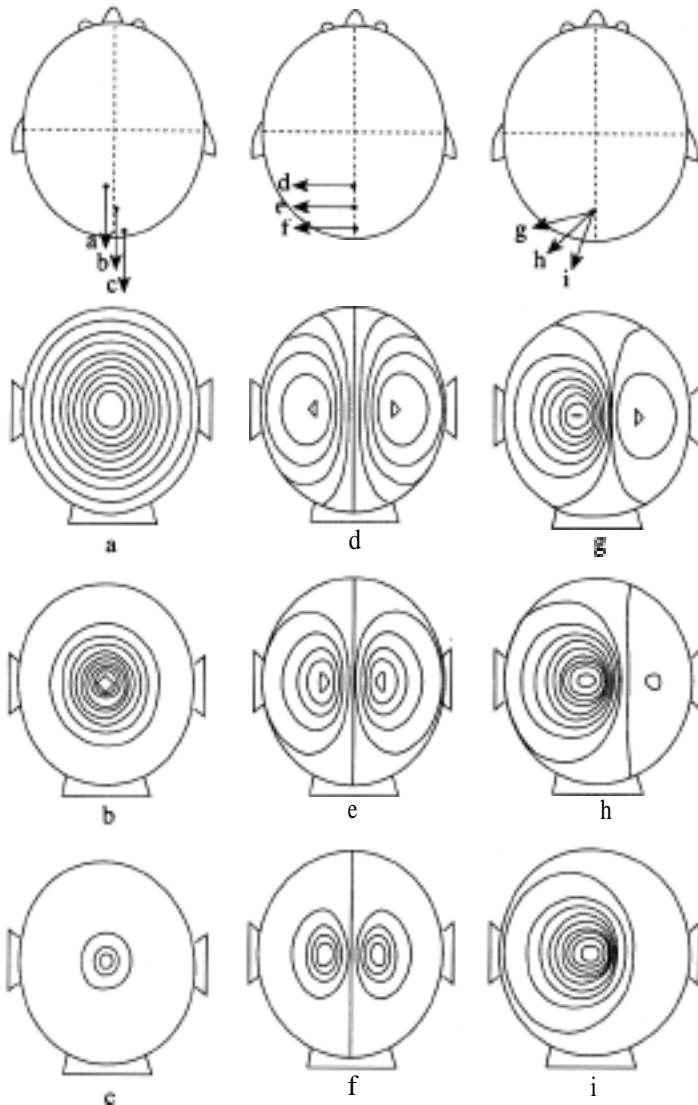
Fig. 12.7 Hypothetical scalp topographic maps, seen from the inion, for the dipoles shown at the top of each column, (From Fender [1987]. Copyright © 1987 Elsevier, reprinted by permission.)

conductivity within one shell (skin, skull, or brain) is constant. But the surfaces defining the shells are broken up into many surface elements. These non-spherical surfaces can be extracted from MRI scans. Boundary element methods are more computationally tractable than three-dimensional finite element methods. While the accuracy of the forward calculation depends on the realism of these models, the important point is that it can, in principle, be solved.

12.7.3 The inverse problem

By contrast, the reverse of the forward problem, namely the inverse problem, cannot be solved in principle. The inverse problem (of estimating current sources in the brain from surface EEG and MEG measurements) is mathematically ill-posed, which means that there is no unique solution in the most general, unconstrained case. The solution is ambiguous because different source configurations can give rise to identical scalp potential and external magnetic field distributions. There are infinite mathematical solutions that are mathematically correct.

While the general inverse problem is ill-posed, it is possible to arrive at a unique solution by restricting the number of possible neural sources or by adding assumptions about the expected form of a many-dipole solution. These general approaches to the inverse problem are the dipole modelling approach, which assumes a small number of dipoles with unknown locations, and the imaging approach, which assumes a large number of (distributed) sources of known locations. The distribution of sources and sinks within a patch observed at a distance greater than. the linear extent of the patch can be well approximated by a single so-called 'equivalent dipole' located in the middle of the patch. A local current dipole is characterized by an amplitude (strength), a position, and an orientation. Simply rotating the orientation of the source at a fixed location can have a major impact on the scalp topography of its field. Deeper sources are generally associated with less compact patterns (less separation between negativity and positivity) and lower signal amplitudes than are shallower sources. Dipole position is defined relative to a three-dimensional co-ordinate system-for example, a Cartesian co-ordinate system (x, y, z) with the origin at the centre of the sphere serving as a head model. The orientation of a source can be described by two angles-for example, its angle relative to a chosen axis (e.g. X-axis), and its angular elevation from the X-Y plane. Each dipole source thus requires six parameters to characterize its activity at one time point: three corresponding to dipole location and three components representing dipole moment (two for angular relations, one for amplitude).

In the few dipole modelling approaches, a small number of equivalent current dipole sources are combined with a head model, and a non-linear least square minimization technique is used to find the six parameters of each of the dipoles that best generates the observed field distribution (e.g. Scherg 1990; Scherg and Ebersole 1993). Non-linear minimization techniques must be used because surface electrical and magnetic fields are non-linear functions of the source location and angle parameters. In this approach, the forward problem is calculated iteratively as dipole parameter values are adjusted until a minimum is found in the error between the recorded data and the values generated by the model.

While it is possible to achieve excellent fits to recorded data using these techniques, there are several problems in interpreting the resulting solution dipoles. One problem has to do with the order of the model. If one assumes there is only one dipole and there is only one dipole, it is possible to localize it with precision. However, if the assumed number is incorrect (e.g. too low), the solution may be misleading. The second problem has to do with local minima in the non-linear fitting procedure. Different starting

points for the fitting procedure (i.e. different initial dipole locations and orientations) can result in different solutions, each of which minimizes error (in the sense that small changes in any of the dipoles lead to higher error). When there are more than two dipoles, it is not practical to exhaustively search the space of solutions. On the positive side, many source configurations, including extended sheets, can be well approximated by a single dipole. And the number of different spatial patterns observed in a typical ERP experiment (as estimated by spatial principal components, for example) is usually not large, indicating that it is possible to describe the activity accurately with a small number of sources.

An alternative to the few moveable dipoles model is a many fixed dipole approach (De Munck *et al.* 1988). Like the few dipole modelling approach, this method requires a mathematical formulation of the forward problem. Unlike the dipole modelling approach, which assumes a small number of dipoles or other current sources the current modelling approach assumes a surface or volume of a large number of possible current elements whose values are to be estimated. It is necessary to calculate the non-linear forward problem for each assumed current element in the volume only once. The most unconstrained of these reconstructions assume a three-dimensional grid of locations, with three orthogonal dipoles at each location. Determining the amplitudes of these dipoles is a linear problem.

The main problem with this approach to source localization is that the number of locations needed to represent the surface or volume of tissue is typically so large that it results in a severely underdetermined linear problem; that is we are faced with a system of linear equations where the number of unknowns far exceeds the number of measurements. Recording data from more electrodes distributed evenly across the volume surface is thus critical for providing a good estimate of the field for which the sources are sought. However, with no additional constraints, it is not possible to recover more dipoles than the number of electrodes divided by three (Fernandez *et al.* 1995).

One way of helping with the underdetermination problem is to use our knowledge of the sources of electrophysiological currents to constrain the reconstruction volume. Restricting the dipolar sources to the cortex, for example, reduces the number of possible locations in x, y, z but still requires solution for the three amplitudes of the three orthogonal dipoles at each of these locations. Additional constraints are provided by arguments based on architecture of the cortex and on invasive recordings suggesting that the primary axis of current flow is normal to the local cortical surface. With dipole locations and orientations fixed, it is possible to estimate as many dipoles as there are recording electrodes. Unfortunately, it takes on the order of 10 000 dipoles to adequately 'tile' the cortex. The resulting underdetermined inverse problem is typically solved by assuming that both the noise at each sensor and the strength of each dipole are independent and of equal variance (which is equivalent to the so-called minimum norm solution) (Dale and Sereno 1993). This solution has well understood properties when applied to the cortex (e.g. it tends to underestimate the depth of sources, but more accurately recovers their tangential location).

Functional MRI could potentially be used to further constrain the inverse problem. Functional MRI has superior and more uniform spatial resolution than EEG/MEG, but its temporal resolution is at least three orders of magnitude coarser. However, using fMRI to constrain EEG/MEG only makes sense to the extent that they measure the same thing. One way to test this is to present stimulus sets that are as similar as possible (given the constraints of the traditional fMRI block design versus event-related designs with EEG and MEG)' and then determine to what extent the independent results are similar to each other. Preliminary studies of this kind suggest that these very different techniques appear to detect activity in similar spatial locations (Barinaga 1997).

To achieve the best combination of spatial and temporal resolution, the techniques could be combined to provide mutual formal constraints upon each other. For example, a weighted linear inverse solution could be calculated where cortical locations with significant fMRI activity would be given a higher prior probability of being EEG or MEG sources (but without completely preventing sites inactive in fMRI images from contributing to the solution) (Dale *et al.* 1996). With this technique, it is essentially possible to extract the time course of activity of an fMRI site, improving upon the lower intrinsic spatial resolution of EEG and/or MEG. It is worth noting that this combination technique is not capable of specifically assigning activity to sources detected with fMRI that are close enough to each other so as to be indistinguishable with MEG and or EEG alone (e.g. two nearby visual areas). Nevertheless, the technique is excellently suited to detecting and localizing overlapping time courses of activity in motor, occipital, parietal, and temporal cortices-which are far enough apart.

12.8 The theoretical power of increased spatial resolution

Improved technology and improved models for looking at data, therefore, have combined to increase significantly the spatial resolution of electrophysiological techniques-both for voltage over the scalp and for determining the underlying spatial distribution of sources. But how much theoretical power does this increased spatial resolution afford? It is important when faced with new or improved technology to remember that increased spatial and temporal resolution does not, in and of itself, provide one with more inferential power. Only if the question asked dictates high spatial resolution does high spatial resolution become an advantage; if a question can be answered using a single electrode, it is easier and more economical by far to use a single electrode rather than a hundred or more. Furthermore, as our ability to resolve differences becomes better, we must be even more cautious about assigning meaning to any difference.

As stated earlier, both cognitive and neurobiological processes are structured at multiple levels. One can define differences at all of these levels-different processes within a given neuron, different neurons, different groups of neurons with or without different connectivity patterns, different areas, different hemispheres, different brains ... all acting on the same or different inputs to create the same or different

outputs, also at multiple levels. Differences at one level are not necessarily indicative of differences at another level, and differences in neurobiological processing are not necessarily indicative of differences in cognitive processing (and vice versa). This becomes especially clear if we examine simple examples from neurobiology.

The visual system is divided across the two cerebral hemispheres such that input in the right visual field influences processing in the left primary visual cortex via the left lateral geniculate nucleus (LGN), and input in the left visual field influences processing in the right visual cortex via the right LGN. No gross anatomical or physiological differences have been reported between the hemispheres for this early stage of perceptual processing; both at a macrostructural and at a microstructural level, striate cortex in the two hemispheres seems to be equivalent in its processing of visual input, Most would agree that the processes going on in the two hemispheres at this level, then, are 'the same' at some important level. Obviously, different neurons are involved, but those neurons seem to be connected the same way and they operate on the same type of input to create the same type of output (cf. other sensory and motor areas). One would not want to say that different 'mechanisms' are responsible for processing right and left hemifield stimuli just because these stimuli create different patterns of neural activity. Yet, clearly, the spatial pattern of electrical activity that will be picked up at the scalp for the same stimuli flashed to the right and the left visual fields will be substantially different. Most investigators would not misinterpret this difference, but that is because we are aware of the physiological and anatomical underpinnings of the process being measured in this case. Without that knowledge, we would not be certain of whether the different pattern we record reflects a theoretically significant difference.

So different patterns of neural activity-patterns arising from different neurons and even different ensembles-do not always reflect true processing differences from a cognitive point of view. Even processing in different brain areas may be psychologically uninteresting. Stimulus intensity, for example, seems to have similar effects across a variety of sensory areas-increased brightness, loudness, or pressure all create similar changes in the neural activity evoked in their respective primary sensory areas, namely an increase in firing rate. The neurons, neural ensembles, and even areas involved are different; however, at another level there seems to be something remarkably similar about the coding of sensory intensity across modalities.

When do differences become theoretically important, then? In the absence of theory, this can be quite difficult to determine. Take the case of determining stimulus location in the visual and the auditory modalities. In the visual modality, stimulus location is represented directly; contiguous points in visual space create activity in contiguous points of the retina and this mapping is maintained across early cortical visual processing areas. Determining a visual stimulus' spatial location, therefore, is simply a matter of 'looking' at its position on the visual map. Deriving stimulus location from auditory signals, however, not so easy. The cochlea orders information by frequency rather than spatial location. Stimulus location information must be derived from the input via comparisons of intensity and phase between the two ears; this information only emerges after several stages of neural processing. Clearly there

is an important mechanistic difference between visual and auditory localization—neurons in different areas with different patterns of connectivity process very different types of information in computationally distinct ways in order to obtain the same conclusion (and possibly to elicit the same motor programmes, such as pointing). Yet even in this case, there is a level at which these two processes are the same-they are both involved in stimulus localization.

These examples serve to illustrate that using spatial information alone-differences in distributions or in neural sources-to answer cognitive questions or to decide between competing psychological or linguistic theories is a difficult endeavour. What a difference *means* depends on the question being asked,, the task(s) used, the measure and its sensitivities, and a clear understanding of the levels of both psychology and neurobiology that are relevant. Especially because any purely spatial distribution is a slice through a process distributed in time, interpreting an observed spatial difference as a meaningful neurobiological, let alone functional/cognitive, difference is problematic regardless of a measure's resolution. However, the increased spatial resolution provided by new methodologies and new analysis techniques does open up new ways of using spatial information to aid our understanding of language processing. Instead of simply looking for a difference-any difference-we may begin to be able to examine how and how much two processes or representations are similar or different. We may, for example, begin creating 'similarity spaces' for the spatial distributions we measure during language tasks and from this gain a picture of the extent to which different cognitive processes use the same neurobiological resources or the extent to which different cognitive processes can be mediated by the same neural connections. More importantly, increasing the spatial resolution of techniques that already have good temporal resolution (and vice versa) increases our ability to resolve the real *spatio-temporal* signal that forms the basis of cognitive processing. In order to truly understand the processes underlying language, we need to know both when and where information becomes available in the brain, and our theories need to recognize the essential links between space and time for neural information processing. Ultimately, then, an increase in our ability to resolve spatial distributions is most exciting not because it gives us more power to test our current theories but because it provides us with an increased ability to create and test new theories that explore language in neurobiologicallymeaningfulterms.

Acknowledgements