High-frequency neural activity and human cognition: Past, present and possible future of intracranial EEG research

Jean-Philippe Lachaux a,b,*, Nikolai Avmacher c,d, Florian Mormann d, Eric Halgren e,f, Nathan E. Crone g

a Lyon Neuroscience Research Center, INSERM U1028, CNRS UMR5292, Brain Dynamics and Cognition Team, F-69500 Lyon-Bron, France
b Université Claude Bernard Lyon 1, France
c Department of Epileptology, University of Bonn, Sigmund-Freud Strasse 25, 53105 Bonn, Germany
d German Center for Neurodegenerative Diseases (DZNE), Ludwig-Erhard-Allee 2, 53175 Bonn, Germany
e Multimodal Imaging Laboratory, Department of Radiology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92037-0841, USA
f Siemens Healthcare, USA

1. Historical background: from Berger to ‘GBR’

1.1. Intracranial EEG and human cognition

1.2. The gamma ‘buzz’

1.3. High-frequency activity or high frequency oscillations?

1.4. Away from oscillations … and back?

1.5. HFA and ERPs

1.5.1. HFA and ERPs are generated by different mechanisms

1.5.2. HFA can distinguish between neural activity increases and decreases, while ERPs cannot

1.6. A unique window into the human brain

Abbreviations: EEG, electroencephalography; iEEG, intracranial electroencephalography; ERP, event-related potential; LFP, local field potential; HFA, high-frequency activity; HFS, high-frequency activity suppression; HFO, high-frequency oscillation; MUA, multi-unit activity; ECS, electrocortical stimulations; ESM, electrocortical stimulation mapping.

* Corresponding author at: INSERM U1028, CNRS UMR5292, 69675 Bron Cedex, France. Tel.: +33 4 72 13 89 00; fax: +33 4 72 13 89 01.
E-mail address: Jp.lachaux@insERM.fr (J.-P. Lachaux)

0301-0082/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.pneurobio.2012.06.008
2. Tell us something new: what can we learn from iEEG HFA about the functional dynamics of the human brain? ........................................................ 283
   2.1. Memory ................................................................................. 284
      2.1.1. Working memory .......................................................... 284
      2.1.2. Long-term memory formation ....................................... 284
      2.1.3. Memory consolidation .................................................. 284
   2.2. Language .............................................................................. 287
      2.2.1. Language production and perception ......................... 287
      2.2.2. Reading ....................................................................... 287
   2.3. High-frequency suppression in the default-mode network .... 288
      2.4. Time, yes, but more than just time ................................. 289
         2.4.1. Time reveals function .............................................. 289
         2.4.2. Time also reveals functional integration and interactions 289
   3. iEEG and cognitive HFA: guidelines for experimental research ... 291
      3.1. Fragments of a lost tale. How many patients make a study? 291
      3.2. Individual-level or group-level statistics? ....................... 292
      3.3. Is an epileptic human brain a good ‘model’ of the healthy human brain? ............................................................. 292
      3.4. Additional technical issues .............................................. 292
         3.4.1. A noisy environment .............................................. 292
         3.4.2. Fast, but not too soon .............................................. 292
   4. Use of HFA to anticipate and prevent post-operative cognitive deficits 293
      4.1. Fast and systematic mapping of major functional networks with HFA: the Grenoble experience ............................................. 293
      4.2. BrainTV: cognitive-field mapping in the human brain ....... 293
      4.3. Participating vs. essential sites in cortical networks ........... 294
   5. Expected conceptual, methodological and technological evolution of iEEG research on cognitive HFA ................................................................. 295
      5.1. Beyond epilepsy: extension to deep-brain stimulations patients 295
      5.2. Novel analysis strategies of iEEG signals ......................... 295
      5.3. Multimodal imaging: hold your hand out! ........................ 295
      5.4. Multi-scale iEEG: micro-, macro-, and multimodal electrodes 296
      Acknowledgements .................................................................. 297
      References .................................................................................. 297

1. Historical background: from Berger to ‘GBR’

1.1. Intracranial EEG and human cognition

In the last few years, an increasing number of research groups have turned to intracranial EEG (iEEG) recordings to study the neural bases of human cognition. iEEG itself is an old technique that dates back to the early pioneers of electrophysiological recordings (Jasper and Carmichael, 1935), soon after Berger’s first application of the EEG to humans. Because of its invasive nature, its use in humans has mainly been restricted to the clinical circumstances of patients undergoing resective surgery for medically refractory epilepsy, in whom it has been used most often to map the cortical networks responsible for seizures. However, iEEG also provides a unique window into the spatio-temporal dynamics of the human brain at work, with submillimeter and millisecond precision, sometimes at the individual neuron level (Engel et al., 2005). This is why scientists have used iEEG to study human cognition since it first came into clinical use. This clinical research has by its nature required multi-disciplinary collaborations between epileptologists, neurophysiologists, neuro-radiologists, and cognitive neuroscientists. The number of such collaborations remained relatively stable for decades, until a shift in scientific focus occurred during the 1990s that boosted the entire field. While early iEEG research was primarily focused on event-related potentials or peri-stimulus time histograms (e.g. Allison et al., 1994; Halgren et al., 1978a,b, 1980; see Lachaux et al., 2003 for review), the emerging trend in ‘cognitive’ iEEG research has been characterized by its strong emphasis on high-frequency neural activity (HFA: 40 Hz and above), which can sometimes be seen in raw iEEG signals (Fig. 1) (Jerbi et al., 2009; Crone et al., 2011; Jacobs and Kahana, 2010). A few early iEEG studies had reported that motor tasks and sensory stimulation could modulate HFA (e.g. Sem-Jacobsen et al., 1956; Chatrjan et al., 1960; Brindley and Craggs, 1972; Halgren et al., 1977), but the field really took off with the ‘gamma buzz’: the marked interest of the neuroscience community in gamma-band synchronization and its role in neural representation and communication fueled by recent experimental and technological advances.

1.2. The gamma ‘buzz’

This phenomenon was initiated by a series of electrophysiological studies in the cat visual cortex (Eckhorn et al., 1988; Gray

![Fig. 1. Task-induced high-frequency activity can be seen in raw iEEG signals. A picture (face stimulus) was flashed foveally for 200 ms while the patient fixated the center of a computer screen. The signal is a raw bipolar recording from the fusiform gyrus (expressed in % of the maximal amplitude value during the time window of interest). HFA is clearly visible in the raw trace (square box), together with the event-related potential evoked by the stimulus, and alpha and beta oscillations (respectively before and after stimulus presentation).](image-url)
et al., 1989), suggesting that object perception might be mediated by rhythmic and synchronous neural firing in area 17 at frequencies around 40 Hz (the ‘gamma’ band), in line with previous theoretical and experimental work (Milner, 1974; Freeman, 1975). This idea was summarized by the dynamic binding or binding-by-synchrony hypothesis (Singer, 1999). As the binding-by-synchrony hypothesis started to receive experimental support in monkeys (Kreiter et al., 1995; Muller et al., 1996; Tallon-Baudry et al., 1996). It was clear that due to its limited spatial resolution EEG could not reveal synchrony between pairs of neurons, but the bet was that the emergence of synchronous neural assemblies would produce macro-scale oscillations, which could be detected in EEG signals with state-of-the-art time-frequency analysis. In fact, such oscillations had already been reported during movement execution in EEG signals (Pfurtscheller et al., 1993; Salenius et al., 1996, and see Pfurtscheller and Lopes da Silva, 1999). In the second half of the nineties, Bertrand and Tallon-Baudry in Lyon published a series of seminal scalp-EEG studies describing energy increases in the gamma-band in response to visual stimuli (summarized in Tallon-Baudry and Bertrand, 1999). Their first contribution was to distinguish between evoked and induced gamma-band responses (or ‘GBR’), which are phase-locked and non-phase-locked, respectively, to the stimulus onset (Tallon-Baudry et al., 1996). Together with a couple of other EEG groups (Lutzenberger et al., 1995; Muller et al., 1996), their focus shifted rapidly to induced GBRs, which were particularly strong when stimuli required feature-binding, as predicted by the binding-by-synchrony hypothesis (Tallon-Baudry and Bertrand, 1999 for review). However, scalp EEG could not reveal the anatomical origin of induced GBR, which immediately motivated the team led by Varela in Paris to replicate the first paradigm of the Lyon group with intracranial EEG recordings from the visual cortex (Lachaux et al., 2000). The study used a grid of electrodes covering the temporo-parietal junction and revealed sensory gamma-band responses in iEEG recordings at frequencies up to 80 Hz (the upper limit of valid investigation for that specific recording system). This indicated that induced GBRs involved in fact a higher and broader frequency range than is visible in scalp EEG, and that were not homogeneous responses of visual cortex, but a collection of responses with distinct latencies and exquisite anatomical specificity. This result echoed an earlier study by Klop et al. showing a power increase in the fusiform gyrus up to 45 Hz, in response to faces (Klop et al., 1999). In the sensorimotor cortex, Crone et al. (1998) had also found increased gamma-band activity during movement execution. This study was not only the first report of task-related HFA in human iEEG recordings, but also provided clear evidence that the spatial distribution of increases in HFA in sensorimotor cortex was specific to the movement of different body parts. It was quickly followed by other iEEG studies showing increased HFA in sensorimotor cortex during visuo-motor behaviors (Aoki et al., 1999; Ohara et al., 2000; Pfurtscheller et al., 2003).

1.3. High-frequency activity or high frequency oscillations?

These pioneering studies launched a steadily rising field demonstrating that cognitive HFAs are ubiquitous in the human brain, and yet exquisitely task-specific. In less than ten years, it has become clear that HFA is an index of cortical processing not only in sensory and motor cortices, but probably in any cortical region that is involved by a task (Jerbi et al., 2009; Crone et al., 2011; Jacobs and Kahana, 2010). This evolving experimental realization of HFA as a general purpose index of cortical processing has widely been considered to be consistent with theoretical extensions of the mechanistic role of gamma-band synchronization, from sensory and motor integration to neural communication in general (Fries, 2005). Nevertheless, human iEEG studies of HFA have also contributed new and provocative findings that have yet to be fully integrated into existing theoretical frameworks.

One burning question for instance is whether cognitive HFA visible in iEEG recordings actually relate to the original phenomenon described by Singer and colleagues. So far, the vast majority of iEEG studies on task-induced HFA have reported broadband energy increases in a frequency band ranging from 40 to 150 Hz, typically. Although such increases have often been called ‘gamma-band responses’, because their frequency extent includes the gamma-band, it appears to be unlikely that they actually reflect the classic, narrow-band, gamma-band synchronization mechanisms hypothesized to subserve, for instance, feature-binding in the visual cortex (Singer, 1999). This is quite ironic, considering that HFA research on HFA entirely emerged from that theoretical framework.

From a mathematical point of view, such a broadband energy increase cannot correspond to a single high-frequency oscillation, or HFO, as it does not possess a well-defined frequency and phase. Instead, Muller et al. (2009a,b,c) have suggested that task-induced HFA might in fact result from an increase of the global spike rate of the underlying neural population (together with asynchronous post-synaptic currents), which would result in a broadband energy increase in bandpass-filtered iEEG signals. This claim has received recent support from studies showing a strong correlation between multi-unit activity and HFA in animals (e.g. Ray et al., 2008; Ray and Maunsell, 2011) and in humans (Manning et al., 2009). Indeed, much of the power of the action potential as recorded extracellularly is in the 50–150 Hz HFA window (Pettersen and Einevoll, 2008). However, this might not be the whole story, since this frequency range also includes synaptic activity, and in a modeling study where these parameters were adjusted to match those experimentally observed, activity between 50 and 150 Hz arose mainly from synaptic currents (Pettersen et al., 2008). In fact, the relative contributions of synaptic and action potentials to the HFAs are influenced by the number of cells that are synaptically activated, the proportion of those that are driven to fire, the locations of the synapses, and the morphology of the neurons, among other factors. Thus, although HFA is highly correlated with multi-unit activity (MUA) (Manning et al., 2009), it is probably not solely generated by MUA. Further biophysical modeling studies that match human anatomy and physiology in cognitive paradigms are needed to confirm that claim.

An important implication is that HFA is reflected in the BOLD signal of fMRI (Mukamel et al., 2005; Nir et al., 2007), because both are fairly non-specific indicators of local neuronal activation. However, BOLD has a significant delay (1–2 s) and smears activity over long periods (~10–20 s) whereas HFA has no delay or smearing, and thus can provide the temporal resolution needed to probe the stages of cognition.

1.4. Away from oscillations … and back?

It should be clear that HFA might correlate with MUA and yet be generated by neural oscillations. The two hypothesis are not mutually exclusive: in a scenario where neural populations would systematically produce rhythmic discharges of action potentials when activated, the production of neural oscillations would always require an increase of MUA. The main argument against an oscillatory nature of HFA is not the strong correlation between HFA and MUA, but the broad frequency extent of HFA. But then, it should be noted that broadband HFA is mostly visible in average
time–frequency representations of neural responses to multiple stimuli, or movements. One possibility, which still awaits confirmation in multiple patients, cortical structures, and experimental situations, is that single-trial responses are characterized by a juxtaposition of several transient components with well-defined frequencies (we can call them elementary time–frequency responses). Elementary time–frequency responses might correspond to local high-frequency oscillations generated by several local neural ensembles, within the larger population recorded by each iEEG electrode (Fig. 2), see (Crone et al., 2011). The broadband HFA increase revealed when averaging across trials would then correspond to the statistical distribution of those elementary responses in the time–frequency domain. The broadband appearance of the average of many such responses might arise from the fact that they are generated by small neural ensembles with different task and stimulus specificity (Gaona et al., 2011). For instance, the precise frequency of gamma-band synchronizations in the visual cortex has been shown to depend on stimulus contrast, not to mention alertness and vigilance level (Ray and Maunsell, 2010; Lakatos et al., 2004). This means that in the visual cortex for instance, two neighboring neural populations with different receptive fields, and processing different parts of a complex image with inhomogeneous local contrasts, would generate oscillations at different frequencies. Therefore, we should not expect all neuronal assemblies recorded by a single iEEG electrode to synchronize at the same frequency. Rather, iEEG is likely to record the sum of several locally generated HFOs, whose frequency and latency vary from trial to trial as stimuli, processing latencies and durations, altertness and vigilance levels also vary. The average response across trials might then cover a broad frequency range, as it is most frequently observed.

To conclude, the broad frequency extent of task-induced HFA may be explained both by a simple increase of the firing rate, and by local synchronization mechanisms producing HFOs at varying frequencies and latencies. Once again, these two scenarios are not mutually exclusive, and a final dissociation between these might require the combination of iEEG and extensive single-unit recordings of underlying neural populations. Until the exact mechanisms generating task-induced HFA are revealed, we should keep in mind their most notable property: they reveal, with excellent time resolution, the participation of local neural populations in the task-at-hand. For this review, we shall therefore refer to the phenomena we observe in iEEG recordings as cognitive ‘HFA’—not ‘HFO’. And we make the strong suggestion that in any future iEEG study on HFA, authors should provide the precise frequency range of the effect, such as HFA “between 40 Hz and 150 Hz”, or more conveniently when the term must be repeated: HFA(40–150). We also propose that authors refer to high-frequency neural activity as “neural oscillations”, only when they have evidence that the activity is rhythmic, or a collection of several rhythmic processes. In the absence of this evidence, the interpretation and discussion of results should acknowledge whether they are made within a conceptual framework that does or does not assume neural oscillations.

1.5. HFA and ERPs

Electrophysiologists might wonder about the relationship between HFA and more ‘classic’ electrophysiological measures used for decades to understand human cognition. In particular, non-invasive studies of cortical information processing in humans have generally used averaged EEG, event-related potentials (ERPs), or their magnetic counterparts, event-related fields (ERFs). These both arise from neuronal currents that are measured within the cortex as local field potentials (LFPs). At the scalp level, ERPs arise from the propagation of the currents through the various tissues of the head whereas ERPs are directly generated by the intracellular currents. Therefore, ERPs are more smeared by CSF and skull, and reflect the activity of more cortical patches than ERPs, but their local generators are two symmetrical limbs of the same current loop (Cohen and Halgren, 2009). In iEEG recordings, intracranial ERPs (iERPs) can be formed from event-related averaging of LFP. Commonly, iERP and HFA are calculated from the same LFPs, and, commonly, they have a similar distribution and time-course but with clear differences (Vidal et al., 2010; Engell et al., 2012), for reasons that we now discuss.

![Fig. 2](image-url) Conceptual schematic of how iEEG broadband gamma activity increases could result from modulation of band-limited gamma oscillations. Area of cerebral cortex is recorded by array of surface electrodes on the left. The power spectrum for one iEEG site recorded under resting vs. active conditions is schematized on the right (arrow). The field of view of each recording site (inverted cones) includes many neuronal assemblies (cylinders) with different functional response sensitivities (larger size and smaller number for illustration). Neuronal assemblies activated by a task (colored cylinders) are synchronized and generate membrane potential oscillations at different band-limited gamma frequencies that depend on the resonant properties of each assembly. Membrane potential oscillations with different center frequencies (colored bands on right corresponding to cylinders to left) collectively contribute to signals recorded by iEEG electrodes and their summation gives a broadband shape to gamma responses in the power spectrum. Note that many more assemblies and bands than can be represented here would be needed to produce the shape of commonly observed spectral responses. Also, different sets of neural assemblies and bands may be involved during different trials of the same task, depending on stimulus properties and a variety of initial conditions. Assemblies not immediately engaged in task-related cortical processing are represented by black (and white) cylinders. Task-related power suppression in alpha/ beta frequencies (red circular area on left and red bands on right) occurs in a wider area and theoretically reflects thalamocortical gating mechanisms that permit or facilitate cortical processing in assemblies with similar but distinct response sensitivities. Reproduced with permission from Elsevier.
1.5.1. HFA and ERPs are generated by different mechanisms

The reasons for the differences between iERP and HFA can be appreciated by considering their contrasting mechanisms of generation. Except for very short-latency primary sensory components, the frequency content of iERP is below 50 Hz, mainly between 1 and 10 Hz, and thus is non-overlapping with HFA. The power of the LFP in the iERP range is also 100–1000 times greater (per Hz) than in the HFA range (e.g., Miller et al., 2009a, b, c) for both temporal and spatial reasons. Temporally, the short synaptic events underlying HFA will not be synchronized unless their timing is synchronized to within a few milliseconds, and asynchronous potentials will tend to cancel each other when they superimpose at the recording electrode. In contrast, the longer synaptic events underlying iERP will superimpose and sum linearly at the electrode even when their timing is off by tens of milliseconds. Spatially, inward synaptic currents in the HFA frequencies return back to the extracellular space over a shorter distance, and thus produce smaller LFP fields than to the longer synaptic events underlying iERP (Linden et al., 2010). Combining these effects, synchronous long-duration synaptic currents can take advantage of the long parallel apical dendrites of pyramidal cells, and the laminar termination of afferents onto these dendrites, to generate LFPs which can summate over relatively long distances, and thus produce the large amplitudes that characterize iERPs (Linden et al., 2011).

1.5.2. HFA can distinguish between neural activity increases and decreases, while ERPs cannot

The special circumstances that lead to large amplitude LFPs in the iERP range imply that iERPs selectively reflect synaptic activity that is relatively long duration, temporally synchronous, and spatially organized onto the apical dendrites of pyramidal cells. Such conditions can arise from either excitatory or inhibitory post-synaptic currents. Consequently, increased iERP can reflect either increased excitation or increased inhibition, or even reflect a more complex modulation. Indeed, a recent study suggests that in hippocampal slices single spikes of inhibitory interneurons terminating on pyramidal cells may be sufficient to produce LFP, but spikes of pyramidal cells themselves are not (Bazet et al., 2010). Other combined modeling and recording studies in hippocampal slices suggest that voltage-gated transmembrane neuronal currents are the major contributors to stimulation-evoked low frequency iERPs (Murakami et al., 2002, 2003). Further, quantitative modeling has found that simultaneous excitatory inward currents at the distal apical dendrites combined with inhibitory outward currents at the soma were needed to reproduce empirical iERP measures (Pettersen et al., 2008). These imply that positive and negative iERP can both reflect either excitatory or inhibitory currents, and that these currents can be either ligand-gated at synapses, or voltage-gated throughout the dendritic surface.

A further limitation of iERP is that their polarity depends on the location of the electrode contact with respect to the generating sources and sinks, which means that the same neural phenomenon can be recorded either as a positive or a negative iERP. Furthermore, because scalp ERPs are the summated propagation of iERPs from different cortical patches, an increased ERP could result from the removal of a canceling iERP (Lutkenhoner, 2003). Current source density (CSD) analysis with MUA can resolve these ambiguities (Einevoll et al., 2007; Ulbert et al., 2001). However, in general, the only reliable inferences that can be drawn from ERPs or iERPs concern the timing and relative amplitude of the underlying generators. But even then, one should keep in mind that ERPs rely on phase-locking of signals across trials and are quite sensitive to jitter in both the phase and latency of neural responses. This means that the amplitude of the ERP might be reduced in an experimental condition relative to another one if its phase varies more across trials, even if its amplitude remains the same.

In contrast to the iERP, power changes in HFA are typically more robust, sometimes detectable even in single trials, and are generally less sensitive to small variations in the timing of behavioral and neural responses. In addition, HFA fluctuations can be correlated in real-time with purely endogeneous processes such as attention fluctuations, mental imagery or spontaneous thought processes (see Section 4.2), while iERP can only be defined in relation to repeated sensory or motor events. As a relatively direct and nonspecific measure of the local high-frequency synaptic and/or unit activity, HFA may be interpreted as an estimate of the intensity of any given population’s engagement in task processing. Thus, HFA is particularly well suited to visualizing the onset, magnitude, and duration of changes in neuronal activity during cognitive tasks, or in relation to endogeneous cognitive processes. In particular, HFA – unlike iERP – can reveal task-induced suppressions of neural activity, which makes them ideally suited for the study of the default-mode network, for instance (see Section 2.3).

1.6. A unique window into the human brain

For all the reasons stated above, we would like to promote the claim that studying HFA with intracranial EEG recordings can provide insights into the neural bases of human cognition that cannot be derived as easily from the study of BOLD effects with fMRI or from the analysis of ERPs. iEEG research on cognitive HFA has now diversified and the objective of this article is not to provide a comprehensive review of this lively field, which can be found elsewhere (e.g. Crone et al., 2011; Jerbi et al., 2009). The first objective of this review paper is to support our claim with illustrative examples. The second objective is to propose guidelines for future iEEG research on cognitive HFA. The third objective is to envision future developments of our research field, which will lead, in our opinion, to several major breakthroughs in human cognitive neuroscience.

2. Tell us something new: what can we learn from iEEG HFA about the functional dynamics of the human brain?

We argue that studying HFA with iEEG is relevant for two reasons: first, because it may serve to provide a link between animal research on the cellular mechanisms supporting cognitive functions and the non-invasive investigation of complex cognitive processes in humans. Second, because it allows one to test specific neuroscientific models of human cognitive processes, which often make predictions at a level of spatial and temporal precision that cannot be tested using other methods. These models may be derived from animal research, computational neuroscience, or cognitive psychology. Throughout this section, we will provide examples illustrating how iEEG studies of HFA are uniquely able to help fulfill these promises.

In those examples, we focus first on iEEG studies of HFA related to human memory and language, then on more general issues in cognitive neuroscience such as global cognitive control via neural activation and deactivation, as well as the temporal dynamics of neural activity and communication in large-scale neurocognitive networks. This selection is purely based on our own fields of expertise and should by no means underestimate that HFA is also highly relevant for other cognitive processes such as perception, attention, and consciousness. In other words, the following examples aim only to illustrate the kinds of questions in human cognitive neuroscience for which iEEG HFA analyses may be
uniquely suited to provide some answers. Thus, it is by no means an exhaustive list of the studies that are possible.

2.1. Memory

The first sections are devoted to memory, because several iEEG studies have indicated that neural activity associated with HFA plays a role in memory processes. We will review studies that investigated HFA during working memory (WM) and long-term memory (LTM) tasks, as well as HFA related to memory consolidation. We will first review the experimental results, and then suggest possible mechanisms by which HFA-related neural activity support memory processes.

2.1.1. Working memory

WM has been conceptualized as the ability to maintain a limited amount of information in an immediately accessible form and to perform cognitive operations on these items (Baddeley, 1986). As this process involves maintaining the information for several seconds, it has been linked to persistent changes in neural activity, e.g. a persistent increase in firing rates (Fuster, 1990; Young et al., 1997). For example, several studies have shown sustained increases of gamma band activity during WM maintenance, which correlates with WM load, in various neocortical regions (Tallon-Baudry et al., 2001; Howard et al., 2003; Mainly et al., 2007) as well as in the rhinal cortex (Axmacher et al., 2007) and in hippocampus (Van Vugt et al., 2010). Beta and gamma oscillations may also be phase-locked between widespread frontal, parietal and occipital areas during the rapid storage, modification and retrieval of multiple memoranda in more complicated working memory tasks (Halgren et al., 2002). These studies have provided insights beyond those of previous fMRI studies, because they have allowed distinctions between transient and sustained responses with a higher temporal resolution than would be possible through analyzing the BOLD response alone.

The HFA modulations observed during working memory tasks may be related to repeated reactivations of neural assemblies, which are each synchronized in the gamma frequency range. This is of course in a framework assuming that HFA implies neural oscillations. In this section, we will avoid discussing that aspect further, and focus on the interpretations of the authors of each study (the same policy will apply throughout the text).

According to an influential computational model, maintenance of multiple items depends on such reactivations during several consecutive cycles of lower-frequency (e.g. theta, between 4 and 7 Hz) oscillations (Lisman and Idiart, 1995; Jensen and Lisman, 2005), which are independently generated in various brain regions (Raghavachari et al., 2006). Recent studies have provided experimental evidence for this model by showing increased coupling of the amplitudes of HFA to the phase of theta frequency oscillations during a continuous word recognition memory task (amplitude-modulated activity at various frequencies between 12 and 46 Hz depending on task condition; Mormann et al., 2005) and, more importantly, during WM maintenance, as compared to inter-trial intervals (amplitude-modulated activity at around 28 Hz; Axmacher et al., 2010a) (Fig. 3). While these observations have been made using intracranial EEG recordings from the hippocampus of epilepsy patients, similar results have been obtained using pattern classification analyses of human MEG data, suggesting that spontaneous re-occurrence of category-specific neural activity is linked to the phases of ongoing theta activity (Fuentemilla et al., 2010; Poch et al., 2011). On the other hand, recordings from the prefrontal cortex of monkeys have shown that neuronal activity representing different items occurs at different phases of low gamma band activity at 32 Hz, not at different theta phases (Siegel et al., 2009). Again, these investigations into the complex spatio-temporal pattern of working memory maintenance cannot be performed using functional MRI alone.

2.1.2. Long-term memory formation

In addition to these observations related to WM, several studies have also demonstrated an increase in HFA during LTM formation. As most intracranial EEG studies have focused on declarative long-term memory processes, we will concentrate on the role of HFA for this type of LTM. In a series of studies, Sederberg and colleagues investigated the role of hippocampal and neocortical HFA recorded with intracranial and subdural EEG electrodes for subsequent memory to words. In their first study (Sederberg et al., 2003), they found that HFA between 28 and 64 Hz at widespread neocortical sites predicted subsequent memory. Similar results were later found in the hippocampus for activity between 44 and 64 Hz (Sederberg et al., 2007a). Furthermore, similar patterns of both hippocampal and neocortical HFA enhancements (between 44 and 100 Hz) were found during successful encoding and retrieval (Sederberg et al., 2007b), consistent with the contextual reinstatement hypothesis. In addition to these positive relationships between HFA and subsequent memory, Fell et al. (2001) found that hippocampal HFA between 32 and 48 Hz was negatively correlated with subsequent memory. Finally, it was found that rhinal–hippocampal phase synchronization between 32 and 48 Hz predicted subsequent free recall of words (Fell et al., 2001). In a continuous recognition paradigm (Fell et al., 2008), a similar, though less pronounced effect was observed for subsequently recognized words (rhinal–hippocampal phase synchronization increases between 37 and 46 Hz), while the reverse effect was found in a lower range (28–34 Hz).

Mechanistically, these effects may be related, via spike-phase coherence (Jacobs et al., 2007), to a precise alignment of action potentials between these two regions or within the hippocampus. Interestingly, hippocampal spike-phase coherence in the theta range (3–8 Hz) was increased for subsequently remembered items (Rutishauser et al., 2010). Such temporal alignment is a requirement for spike-timing dependent forms of synaptic long-term potentiation (so-called Hebbian plasticity: Axmacher et al., 2006; Fell and Axmacher, 2011). A facilitating mechanism to achieve phase synchronization in the gamma band could be the coupling of otherwise independent theta rhythms (Mormann et al., 2008a) between relevant areas (Fell et al., 2003).

Even more detailed information has been obtained by studies using linear microelectrode arrays with 24 contacts on 150 μm centers, which have explored the flow of information in the anterior temporal lobe during memory. Chan et al. (2011) found that high-gamma activity in the anteroventral temporal lobe distinguishes between words referring to objects vs. those referring to animals. This distinction is present in the first pass, feedforward activation, characterized as an EPSC (i.e., current sink) in layer IV before 200 ms. It could be seen as defining the type and timing of semantic information that may be projected to the hippocampus as the raw materials of its memories. Conversely, in case studies during retrieval (Knake et al., 2007), especially of autobiographical remote memories (Steinworth et al., 2010), sustained increases in HFA were found in the superficial entorhinal cortex layers (neocortical recipient).

2.1.3. Memory consolidation

A group of so-called two-step theories of memory formation have suggested that the initial encoding stage, which leads to the formation of labile memory representations, is followed by a second stage called memory consolidation, which renders memory traces stable against interfering inputs (Marr, 1971; Buzsaki, 1989; McClelland et al., 1995; Squire and Alvarez, 1995) and appears to be linked to very fast (in animals, around 200 Hz)
bursts of “ripple” oscillations (Buzsaki et al., 2002). Evidence for human ripples has been found both with microelectrode (Bragin et al., 1999a,b; Staba et al., 2002; Worrell et al., 2008) and with macroelectrode (Urrestarazu et al., 2007; Axmacher et al., 2008; Worrell et al., 2008) recordings, as well as with semi-invasive foramen ovale electrodes (Clemens et al., 2007). The rate of such ripples within the rhinal cortex predicts memory performance after sleep between individuals (Axmacher et al., 2008) (Fig. 4), further supporting the link between these events and the processes of memory consolidation. These oscillations, which in humans have a lower frequency at around 100 Hz, in the frequency range of interest for this review, should not be confused with so-called “fast ripples” at 500 Hz, which occur predominantly in the vicinity of the epileptic focus and appear to be related to pathological processes (Bragin et al., 1999a,b; Staba et al., 2002; Foffani et al., 2007).

The mechanisms by which sharp wave ripples might promote memory consolidation are starting to be understood. It has been shown that these events are temporally coupled to the occurrence of neocortical sleep spindles (Siapas and Wilson, 1998; Sirola et al., 2003; Clemens et al., 2007), which, in turn, increase after learning (Gais et al., 2002) and may promote neocortical plasticity via an accumulation of intracellular calcium (Sejnowski and Destexhe, 2000). Therefore, sharp wave ripples may be related to the induction of neocortical plasticity processes by the hippocampus. Results from three studies have suggested an even closer link between sharp wave ripples and the cellular basis of memory processes: first, Behrens et al. (2005) demonstrated in vitro that stimulation paradigms typically used for the induction of synaptic plasticity, also increase the rate of sharp wave ripples. Second, King et al. (1999) showed in vivo that the induction of long-term potentiation is facilitated during sharp wave ripples. Finally, it has been shown that high frequency oscillations such as ripples are potent inductors of synaptic long-term potentiation (Yun et al., 2002). Ripple oscillations appear to be temporally linked to reactivation of memory traces: reactivation has been predominantly studied in rodents, where content-specific representations of spatial environments can be investigated using recordings from spatially selective hippocampal place cells (O’Keefe and Dostrovsky, 1971; O’Keefe, 1976). Following exploration of such environments – accompanied by sequential firing in place cells which represent the locations along the rat’s track – these same sequences re-occur spontaneously during sleep (Pavlidis and Winson, 1989; Wilson and McNaughton, 1994; Kudrimoti et al., 1999; Lee and Wilson, 2002), as well as during quiet resting periods (Foster and Wilson, 2006; Diba and Buzsáki, 2007).

It should be noted that a couple of recent studies questioned the classical view of memory consolidation as a one-time process and indicate that even consolidated memories may become vulnerable toward new interfering information. As a result, they need to be strengthened again, a process labeled as “reconsolidation” (Nader and Hardt, 2009). Moreover, it has been shown that consolidation

---

**Fig. 3.** Cross-frequency coupling during multi-item working memory. (A) During maintenance of multiple items (trial-unique novel faces), hippocampal oscillations in the beta/gamma frequency range occur predominantly during a specific phase range of lower-frequency oscillations in the theta range. These data support a computer model of working memory according to which multiple items can be simultaneously maintained by a multiplexing of cycles of high-frequency activity during specific phases of low-frequency oscillations. This effect is more pronounced during the delay period of the working memory task (B) as compared to an inter-trial baseline (C) as well as compared to surrogate data (not shown). Figure modified from Axmacher et al. (2010a).
depends on the possibility to integrate novel information into pre-existing schemata (e.g. Tse et al., 2007; Van Kesteren et al., 2010). However, it is still an open question whether consolidation and reconsolidation depend on separate neurophysiological mechanisms (McKenzie and Eichenbaum, 2011)—possibly, reconsolidation requires similar processes of replay linked to high-frequency ripple oscillations as consolidation does.

Finally, memory consolidation has been linked to sleep. Several studies using linear microelectrodes contributed to understanding how HFA is modulated during NREM sleep, and which layer it is coming from. These experiments recorded wideband activity of a cortical column from the pia to the white matter. Richard Csercsa in Istvan Ulbert’s lab (Csercsa et al., 2010) described the laminar distribution of HFA, MUA and CSD during the slow oscillation in humans. Increased HFA during the upstate is due to a powerful layer II/III sink, and reflects increased synaptic activity in supragranular layers due to increased firing by both infragranular and supragranular pyramidal cells. Using the same techniques, Cash et al. (2009) found that the main (surface negative at ~440 ms) component of the human K-complex is a cortical
downstate, with a profound decrease in synaptic activity and consequently HFA in the same supragranular layers. This could be linked to cognition because the K-complex can be evoked as well as spontaneous.

Taken together, these studies suggest that HFA supports both working and long-term memory processes through different mechanisms: on the one hand, persistent increases of HFA may support the sustained maintenance of information during the delay phase of WM tasks, possibly linked to specific phases of simultaneous theta band activity. On the other hand, transient increases of HFA may be relevant for the initial encoding and subsequent consolidation of memories by facilitating hippocampal spike-timing dependent plasticity as well as neocortical long-term potentiation.

2.2. Language

2.2.1. Language production and perception

Our understanding of the neural mechanisms underlying most perceptual, motor and cognitive processes has greatly benefited from the precision of invasive micro-recordings in animals. Unfortunately, animal studies are of limited use to understand cognitive functions unique to humans. One notorious example is the production and comprehension of elaborated language, in both its oral and written forms. Because of the impossibility of animal studies, our understanding of the neural mechanisms supporting language is not as elaborate as for other major cognitive functions shared with other species, such as memory. However, we do know (Price, 2000), that language is supported by large-scale networks of cortical areas that are widely distributed across occipital, temporal, parietal, and frontal lobes of the dominant hemisphere. Although it is rare for all of these areas to be sampled with iEEG electrodes, it is often possible to sample many of these areas and to measure the magnitude and time course of neural activation at each site. This, in turn, allows researchers to determine whether different sites are activated in sequence, in parallel, or in a cascading fashion. Visualizing the pattern of these fine temporal dynamics is essential for building and testing anatomically constrained models of human language.

Until recently, our knowledge about human language networks derived almost exclusively from clinical lesion studies and from non-invasive neuroimaging studies, using either fMRI, PET, and to a lesser extent, EEG and MEG (Price, 2010; Salmelin, 2007). This non-invasive approach has proved tremendously useful to disassemble language processes into distinct subcomponents and associate them with specific brain regions and stages of each task. However, they do not combine sufficient temporal and spatial resolution for making contact between functional anatomical and psycholinguistic models of language.

For instance, Indefrey and Levelt (2004) performed a meta-analysis of 82 non-invasive studies, to associate specific neuroanatomical regions and time ranges, or stages, in a model of the core representations/processes supporting word production. Though some representations depended on representations at earlier stages, the model predicted that it was not necessary for each stage to be completed before another stage began: initial processing at each stage propagated through the network in a serial fashion, but it triggered a parallel and/or cascading dynamics at the level of large-scale cortical networks. MEG recordings were used to test this model in part with a picture naming task, using phase-locked responses similar to ERPs as a measure of the onset latency of processing at different cortical sites (Levelt et al., 1998). However, MEG data were inconclusive, because it is somewhat difficult to use ERPs to measure and interpret the duration and amount of neural activity (see Section 1.5), which are important for identifying cascading network activations, as well as the relative contribution of simultaneously active areas to overall task processing.

HFA, as measured with iEEG, can serve as a particularly useful index of cortical activation with a temporal precision adequate, in most cases, to contrast and compare the time courses of cortical activations in the various language areas (Canolty et al., 2007; Crane et al., 2001b; Edwards et al., 2010; Pei et al., 2011; Pasley et al., 2012). A variety of language tasks have been studied with that approach, including visual object (picture) naming (Crane et al., 2001a; Edwards et al., 2010), auditory word repetition (Crone et al., 2001b; Flinker et al., 2011; Pei et al., 2011; Towle et al., 2008), sentence comprehension (Brown et al., 2008), and a variety of auditory speech perception tasks (Chang et al., 2010, 2011; Crane et al., 2001a; Edwards et al., 2005; Sinai et al., 2009; Towle et al., 2008), and reading (see next section). In general, these studies have demonstrated spatiotemporal patterns of activation consistent with cascaded processing dynamics. For example, a recent study of verb generation and picture naming (Edwards et al., 2010) demonstrated a serial progression of activation consistent with distinct stages of perception, semantic analysis, and speech production. Nevertheless, there was substantial overlap in the temporal envelopes of HFA increases observed in areas where the onset of activation appeared to be sequential. What is not known, however, is whether the total duration of activation at each site is really necessary for successful task completion, i.e. whether temporally overlapping activations are obligatory and thus whether processing is truly cascaded.

2.2.2. Reading

Among all language functions, reading might be the most difficult one to study with non-invasive techniques. One complicating factor is the speed at which it occurs and its reliance on oculomotor processes: reading a text like this one involves several eye movements per second to scan through sentences and extract their meaning. The temporal resolution of fMRI and PET is too slow to follow the time-course of neural processing of each individual word or word group. And while EEG and MEG are sufficiently fast, their signals mix together contributions from several nodes of the reading network, and most importantly, muscular artifacts produced by eye-movements. In contrast, intracranial EEG largely avoids such pitfalls and appears to be particularly well-suited to studies of such a dynamic, ongoing process as reading.

Of course, this is not to underestimate the contribution of non-invasive imaging. Each of the main subprocesses of reading has been thoroughly investigated in healthy subjects: the visual recognition of written word-forms, grapheme-to-phoneme conversion and semantic and syntactic analysis of sentence components (Price, 2000; Demonet et al., 2005). By comparing brain responses to meaningful vs. meaningless letter strings, fMRI and PET studies have located semantic processes in the inferior frontal cortex and around the angular gyrus, while EEG and MEG studies have associated those processes with a response component around 400 ms after stimulus presentation: the N400 (Devlin et al., 2003; Kutas et al., 2000; Dale et al., 2000). Similar approaches have revealed that word-form recognition activates a small cortical region in the inferior temporal cortex called the Visual Word-Form Area, in direct association with a deflection of EEG and MEG signals 200 ms after visual word stimulus onset (Bentin et al., 1999; Cohen et al., 2000).

Yet, iEEG studies have undeniably revealed the neural mechanisms of reading with unmatched precision. The first intracranial EEG responses to written words were in fact reported more than twenty years ago, some of them at the single-unit level (Smith et al., 1986; Halgren et al., 1994a; Allison et al., 1994; Ojemann et al., 1988). Heit et al. (1988, 1990) showed for instance unit responses in the hippocampus and medial temporal lobe that
were selective for specific words, and were shown to occur during the N400. Slightly later, in a particularly exhaustive study, Halgren et al. (1994a,b) used iEEG to describe a wave of event-related potentials concomitant to single-word reading which propagates between 190 ms and 600 ms through the occipital, temporal, parietal and frontal lobes in regions coinciding with fMRI localizations of the reading network. In later work using multi-electrode arrays, Halgren et al. (2006) demonstrated that the first pass of activity down the ventral stream in the temporal lobe into layer IV is associated with word form encoding, and the N400 is associated with recurrent upper layer activity. Recently Chan et al. (2011) showed that this first-pass layer IV MUA and current sinks in the anteroventral temporal lobe are already selective to the semantic content of words. This first-pass lexical component is continued in Broca’s region by additional iERP components that can be related to grammatical and phonological encoding (Sahin et al., 2009).

In recent years, applications of time–frequency analysis to iEEG signals have shown that the recognition of visually presented words is also concomitant with a cascade of HFA, between 40 and 150 Hz (Crone et al., 2001a; Tanji et al., 2005). In a study emphasizing either semantic, phonological or purely visual analysis of words and letter strings, Mainy et al. (2008) observed HFA between 40 Hz and 150 Hz in the ventral occipito-temporal cortex in relation to word form analysis, in the superior temporal cortex and posterior part of the inferior frontal cortex during phonological analysis, and in the anterior part of the inferior frontal cortex during semantic analysis. HFA was characterized by a posterior–anterior spread between 200 and 400 ms from the left temporal to the left frontal lobe. A second study by the same group showed that HFA in the reading network is also strongly modulated by attention (Jung et al., 2008): in a simple story-reading task contrasting an attended and an unattended story, attended words elicited sustained HFA in the entire network while ignored words only triggered transient HFA in the ventral temporal cortex. This study suggested that word comprehension requires a widely distributed network of HFA including the frontal cortex, which only reacts during attentive reading.

Future studies are expected to build on the aforementioned preliminary observations to study activations and interactions within the reading network with iEEG under more natural conditions (see Vidal et al., 2012 for a first attempt). Eye-tracking devices can precisely monitor gaze position and fixation duration during sentence reading (Rayner, 1998). Such measures have been extensively used in psychophysics, but rarely in relation to neurophysiological processes. Yet, our understanding of the reading network will greatly improve when electrophysiological parameters such as the amplitude, latency, duration, and localization of HFA and correlations between HFA in different brain regions will be related to experimental parameters such as word type (verb, noun, abstract or concrete), context or syntactic complexity.

2.3. High-frequency suppression in the default-mode network

Over the last fifteen years, functional neuroimaging studies have revealed a fronto-parietal network more active metabolically during rest than during attention-demanding tasks (Gusnard and Raichle, 2001; Mazoyer et al., 2001). That network, often called ‘default mode network’ or DMN, includes the posterior cingulate cortex and the precuneus, part of the medial temporal lobe, part of the ventral lateral and medial prefrontal cortex, and the lateral and inferior parietal cortex. It has become evident that any cognitive task requires not only the activation of specific brain regions relevant to the task, but also the deactivation of brain regions irrelevant to the task that might interfere with task-related regions. In tasks that require attention to external stimuli, the ensemble of irrelevant regions is the DMN; and failure to deactivate the DMN during an attention-demanding task actually impairs performance (Weissman et al., 2006).

Despite considerable advances regarding the understanding of the spatial organization of the DMN and its main functional correlates, very little is known about the precise neural mechanisms taking place in that network, the fine-scale temporal dynamics of neural activity suppressions, and their precise relationship with behavior. This is mainly due to a lack of knowledge about the electrophysiological correlates of DMN metabolic activity decreases. BOLD deactivations have often been interpreted as local cortical inhibition phenomena, that is, transient reductions of the mean firing rate of large neural populations. Firm evidence supporting that claim, combining simultaneous BOLD and multi-unit activity recordings in the DMN, is still lacking however. The best evidence so far was established outside the DMN, in the primary visual cortex: Shmuel et al. (2006) reported a tight relationship between negative BOLD responses and neural activity decreases in monkeys during visual stimulation.

Several studies have shown that the local generation of HFA is often simultaneous with an increase of the BOLD signal (Niessing et al., 2005; Lachaux et al., 2007; Ojemann et al., 2010; Engell et al., 2012). If the BOLD signal correlates with HFA strength, and if it diminishes in the DMN during attentive processing, then HFA strength should also diminish in the DMN under similar attention conditions. This predicts the existence of high-frequency activity suppressions (HFS), that is, transient energy suppressions in the gamma-band during the processing of sensory stimuli. Lachaux et al. (2005) reported a first example of HFS, or ‘negative gamma-band response’ outside the DMN: a visual stimulus that was flashed foveally induced a transient energy suppression between 40 and 150 Hz, in iEEG recordings of the primary visual cortex, in the periphery of the retinotopic map. This negative response mirrored the neural deactivation later reported by Shmuel et al. in monkey V1 (Shmuel et al., 2006). It was also the opposite of positive gamma-band responses induced by visual stimuli in the fusiform gyrus of the same patients. Therefore, sensory stimuli induced both energy increases and decreases at high-frequencies, similar to changes observed in lower frequency bands, such as alpha (Rihs et al., 2009).

Later, the same team reported a similar HFS in the ventral lateral prefrontal cortex (VLPCF) in response to written words during a reading task (Lachaux et al., 2008). This time, the authors proposed an explicit link with the DMN based on previous fMRI studies including the VLPCF in the DMN, and based on the observation that HFS in the VLPCF only occurred when participants processed stimuli attentively. This result was consistent with the fact that DMN-negative BOLD responses are enhanced by attention (e.g. Weissman et al., 2006).

Miller et al. (2009a,b,c) later reported a similar HFS in the DMN during active behavior vs. rest. Ossandon et al. (2011) reported a tight relationship between HFS and behavior throughout the DMN during a visual search task: HFS duration matched precisely the duration of the search, on a trial-by-trial basis. However, the onset latency of HFS was not uniform in the DMN, suggesting that the DMN does not react ‘as a whole’: the VLPCF reacts first, quickly followed by mesial structures. This sequence, which could not have been revealed with fMRI, supports the view that the DMN is temporally and functionally fragmented. The next question, still unanswered, is whether the reaction of the VLPCF has a causal influence on the other components of the DMN.

These results, confirmed by recent studies (Jerbi et al., 2010; Dastjerdi et al., 2011) might have strong implications for our understanding of HFA. To date, cognitive HFA has been mostly
thought of as a transient phenomenon, triggered by, or supporting, short cognitive processes. But the existence of HFS implies that DMN neural populations produce HFA continuously except when external stimuli are attended. It is therefore the suppression of HFA that is transient, not their production. Indeed, Ossandon et al. (2011) showed that HFA re-appears as soon as attentive processing stops, within 100 ms or less, as if DMN HFA had to fill any gap between two external demands. Therefore, HFA should not be seen necessarily as a transient phenomenon, at least in the DMN. This further triggers the question of whether pathological HFA, within the epileptogenic network, might also occur continuously as in the DMN and might possibly participate in seizure generation.

2.4. Time, yes, but more than just time

2.4.1. Time reveals function

We now shift away from specific functional systems to illustrate more general applications of iEEG research on HFA. One aspect which always needs to be emphasized is time. iEEG studies in general have revealed the timing of neural processes underlying human cognition with outstanding precision. For decades, the idea of timing, at least in human cognitive neuroscience, has been strongly associated with latencies of well-known event-related potentials, such as the N170 (an activity peak in the fusiform gyrus 170 ms after the presentation of a face stimulus) (Bentin et al., 1996). Time information was mostly used to validate models from experimental psychology, computational neuroscience or animal research concerned with the sequence of mental operations necessary to produce a response to a stimulus. The discovery of task-induced HFA has begun to shift the focus of iEEG studies to different aspects of temporal information. For example, the high signal-to-noise ratio of iEEG makes it possible to visualize task-related modulation of HFA in single-trials and to correlate its latency and duration with behavioral performance on a trial-by-trial basis (Fig. 5). This kind of visualization allows for an immediate functional dissociation between HFAs with different temporal envelopes. During a visual search task for instance, HFA in the fusiform gyrus is short and time-locked to stimulus onset, while HFA in the posterior occipital cortex is sustained throughout the search, indicating the fusiform gyrus does not participate in the search process, while the posterior occipital cortex probably does. fMRI studies lack the sufficient temporal resolution to detect such a functional dissociation based on timing only.

Further, the timing of iEEG HFA often shows that a single cortical region can process information at multiple stages or levels of task performance. For example, Bastin et al. (2012) have studied HFA in the posterior parahippocampal gyrus (PPG) in response to several categories of pictures, including landscapes, faces and words. The PPG generated an initial category-specific HFA to landscapes, at latencies earlier than 200 ms, with no response to other visual categories, in agreement with the fMRI literature (Epstein and Kanwisher, 1998) and parahippocampal latency analyses in humans at the single neuron level (Mormann et al., 2008b). However, a second task contrasting different types of scene analysis (allo-centric vs. ego-centric evaluation of distance) revealed a later HFA, between 200 ms and 500 ms following stimulus onset, which was task-specific. The early HFA response was thus content-specific – it corresponded to identification of the stimulus as a landscape – while the late HFA response was task-specific, i.e. it supported the extraction of task-relevant information from the stimulus. This temporal dissociation echoes a previous study from the same group showing a succession of two HFA increases in the word-form area during attentive and non-attentive reading: HFA first increased in response to words independent of task-condition, and was followed by a second HFA increase contingent upon the task-relevance of the stimulus (Jung et al., 2008).

Another example where the combined timing and spatial resolution of iEEG can help resolve a controversy from fMRI concerns the origin of selective responses in posteroventral occipitotemporal cortex to words representing objects vs. animals. These posterior areas differentially respond with BOLD to pictures of objects vs. animals regardless of task, but to words representing objects vs. animals only when the task seems to require access to the visual representation of the item (Devlin, 2005). This selective BOLD response could reflect a local computation that is enhanced by strategic processes such as attention during such tasks. Alternatively, this response could reflect later top-down projections from semantic areas. Using laminar multielectrode arrays to record HFA, MUA and current sinks in different cortical layers, Chan et al. (2011) showed that first-pass activation of layer IV in the anterodorsal temporal lobe is already differentially responsive to words representing objects vs. animals at 200 ms, suggesting that the differential BOLD response in posteroventral occipitotemporal cortex reflects top-down activation projected back from more anterior areas. Note that both top-down and bottom-up activation should evoke a BOLD response, provided that they increase the level of neuronal activity. Although BOLD effects may have different latencies and characteristics in different cortical layers, these appear to be due to propagation within the microvascular tree rather than major differences in neurovascular coupling (Tian et al., 2010).

2.4.2. Time also reveals functional integration and interactions

A full account of the functional network dynamics during cognitive tasks requires not only a description of the temporal envelopes of cortical activation at each individual recording sites, but also an understanding of how processing is integrated across large-scale cortical networks and how different nodes of those networks interact with one another. For example, if iEEG HFA analysis reveals substantial temporal overlap in the activation of two different regions, can we tell whether processing in these two regions is mutually dependent or whether processing in one region has a substantive impact on processing in the other region? Are the two regions jointly participating in a representational/processing stage or is one propagating the results of its processing to the other region for further processing? These questions are notoriously difficult to answer, and indeed, the conceptual framework for asking them is still evolving. Yet the answers to these questions may have some very practical clinical ramifications. For example, common language tasks such as picture naming require the cooperative activity of large-scale networks of cortical sites, and it may be difficult to determine the functional role, and thus the relative importance, of any single cortical site based on its functional activation alone. When deciding whether to resect or spare such a site, it might be particularly useful to know whether a given site is common to several different task-specific cortical networks or is an important hub in one such network. This knowledge would be particularly useful if the given site is also part of the network of sites responsible for seizure generation or propagation.

These considerations have motivated a growing number of studies using sophisticated signal processing algorithms that leverage the timing information in iEEG signals to uncover higher-order integration and interactions between and within brain regions. To investigate the integration of cortical processing across large-scale cortical networks, many investigators have measured the degree to which gamma oscillations at different sites are synchronized with each other. Using a variety of signal processing approaches, including phase-synchrony coupling analyses (Lachaux et al., 1999; Tallon-Baudry et al., 2001) these
investigations have been grounded in the gamma-band synchronization hypothesis (Varela et al., 2001). To date, this approach has mostly been applied to non-human primate electrophysiology (Gregoriou et al., 2009) and human non-invasive EEG (Rodriguez et al., 1999) to understand parallel processing across widely distributed neuronal assemblies. In iEEG signals, synchrony is mostly visible in lower frequency bands (e.g. around 20 Hz, Tallon-Baudry et al., 2001; Lachaux et al., 2005), but the spatial and temporal precision of intracranial EEG make it particularly well-suited for studying gamma synchronization as well, as already shown by Fell et al. (2001).

That framework is rapidly extending to include increasingly sophisticated approaches. For instance, graph-theoretical measures have been applied to networks synchronized in the gamma frequency range, to characterize both physiological networks, with scalp EEG (Palva et al., 2010; Kitzbichler et al., 2011; Zhou et al., 2012), and pathological networks with iEEG (Wilke et al., 2011). A new measure of multivariate phase-based interactions has also been applied to iEEG data (Canolty et al., 2012), to separate direct and indirect interactions between channels, which is not possible with the traditional measure of phase synchronization (Mormann et al., 2000). Other approaches try to understand how different frequency bands might interact to facilitate neural integration. For example, recent studies have shown that HFA is modulated by theta oscillations in large-scale brain networks (Canolty et al., 2006; Mormann et al., 2005), generating a growing interest in cross-frequency coupling (Canolty and Knight, 2010) as a mechanism for large-scale integration.

Yet another approach has focused on task-related causal influences among network components, often referred to as “effective connectivity” (Bressler and Tognoli, 2006). This framework emphasizes the impact of processing in one cortical site on processing in other cortical sites activated by a task. Because these causal influences are expected to contribute to network function on timescales of tens to hundreds of milliseconds, iEEG is particularly well suited to study them. A recent iEEG study used multivariate autoregressive modeling (Korzeniewska et al., 2008) to investigate the event-related temporal dynamics of directionally specific causal influences during a set of complementary word production tasks (Korzeniewska et al., 2011) (Fig. 6). This study demonstrated robust and rapidly changing causal interactions at high-gamma frequencies among widely distributed iEEG recording sites. Importantly, the location and timing of these interactions depended on contrasting stimulus modalities and response modalities, and were consistent with anatomically constrained models of the functional anatomy and dynamics of word production. To investigate the relationship between activation at individual sites and causal interactions between sites, the number and magnitude of causal influences that each site had on all other sites was integrated over task-relevant time windows. Interestingly, the sites with the most numerous and prominent influences on other sites in the network, were those with the greatest activation, suggesting that these sites were important nodes in the overall task-specific network. Moreover, although only some of these sites were tested with electro-cortical stimulation (ECS) mapping, there was a remarkable agreement at several sites evaluated with both iEEG and ECS.

Finally, Vidal et al. (2012) reasoned that in most networks, components produce, process and transmit information all at the same time. The logical implication is that the amount of cortical computation performed by two interacting neural populations should co-vary in time. If follows that if HFA quantifies local cortical computation, then the time fluctuations of HFA measured in two interacting neural populations should also covary. In other

Fig. 5. Timing of HFA dissociates between four distinct functional roles during visual search. The four panels show single-trial HFA during visual search in four anatomical sites (FEF: frontal eye field; DLPFC: dorso-lateral prefrontal cortex). Color codes band-limited power of iEEG signals in the [50–150 Hz] frequency range, expressed in % increase or decrease relative to the average value across the entire experiment. Stimulus onset occurs at 0 ms, yellow dots indicate reaction time. The task required that patients find a target (tilted letter ‘T’) embedded in an array of distractors (tilted ‘L’s with random orientation). The four sites are characterized by dissimilar response timing, indicative of clearly distinct functional roles.
words, correlation between HFA envelopes should logically measure network interactions, and as such, be modulated by task-demands. They confirmed that hypothesis during a sentence-reading task, showing interactions specific to semantic processing between two major components of the reading network: the middle temporal gyrus and the inferior frontal gyrus.

3. iEEG and cognitive HFA: guidelines for experimental research

If, as we hope, the previous examples are convincing evidence that iEEG research on HFA can provide unique information about the human brain at work, we expect the number of contributors to this field to increase at a fast pace, as it has already started to do in the last few years. This review is a unique opportunity to combine the experience of five research groups with long-standing expertise in iEEG research, and to provide some guidelines for experimental research on HFA. Human intracranial electrophysiology comes with specific constraints, which require some care.

3.1. Fragments of a lost tale. How many patients make a study?

A major constraint of iEEG is its limited sampling of the human brain. Recording techniques vary across epilepsy centers: 2-D cortical grids provide a spatially continuous coverage of extended superficial cortical areas such as the motor cortex, while depth electrodes penetrate deep into the cortex to reach internal structures such as the cingulate cortex or the hippocampus. Yet, in all cases, electrodes record only from a fraction of the total brain volume (on the order of 1%, see Halgren et al., 1998), which means that iEEG never provides a comprehensive view of global brain dynamics, but only ‘fragments of a lost tale’.

To deal with this limitation, most iEEG studies have chosen to focus on a small set of brain structures, for instance the medial temporal lobe, and to report only results that are reproducible across patients. This approach is very much in line with the tradition of animal electrophysiology, which has to deal with the same spatial sampling issue: for instance, most studies performed in monkeys target a very specific structure and report on data recorded in a few animals only. Note that the number of subjects reported in iEEG studies is typically lower than for non-invasive neuroimaging (fMRI) or electrophysiological (scalp EEG or MEG) studies for two reasons: (a) the number of patients undergoing iEEG studies for clinical purposes at any one center is typically small and (b) the high signal-to-noise ratio of iEEG recordings allows statistically significant results to be obtained in single subjects. Robust and stable HFA responses can be measured in single subjects performing a relatively small number of trials of a task, and multiple blocks of the same task can be recorded in the same subject to demonstrate the test–retest reliability of these responses. In these cases, “grand averages” are simply not needed, and at the high spatial resolution of iEEG, may be inappropriate and potentially misleading because of the risk of averaging responses from neuronal populations with different responses to the same task. Under these circumstances, it is only necessary to show that HFA responses with task-relevant spatial and temporal profiles are reproducible across subjects. Because of the potential effects of epilepsy on normal cortical networks and because no two iEEG electrode configurations (or human brains) are identical, the number of subjects needed to demonstrate reproducibility is arguably more than the two-subject criterion used in publishing studies of nonhuman primates, but if the results are compelling it need not be much more.

Other studies favor a more global approach to report on the global brain dynamics, as much as possible. Data are collected in several patients (often more than 10), and the results are organized into regions of interest, which include anatomical sites with robust task-related responses across patients. In practice, any cortical

Fig. 6. Event-related causal (ERC) interactions between iEEG sites with HFA during an auditory word repetition task. Three sequential time intervals of the task are shown: auditory word perception (stimulus to median stimulus duration, left panel), word retrieval and response preparation (between median stimulus offset and median response onset, middle panel), and spoken response (including 750 ms following the median response latency, right panel). Arrows indicate the directions and intensities of statistically significant increases in event-related causal (ERC) interactions between recording sites. The width and color of each arrow both represent linearly the magnitude of its integral ERC flow. Color scale (on left) has the same range for all ERCs, scaled from minimum to maximum (10% of the smallest ERCs not shown). The integral (sum) of ERC outflows is illustrated by semi-transparent purple circles. The radius of each circle is proportional to the normalized sum of statistically significant event-related increases in causal interactions directed outwardly from the site (originating at the site). Plots of event-related iEEG HFA are shown for select iEEG sites (insets). Only statistically significant power increases and decreases are plotted. Stimulus onset is at 0 s. Vertical markers indicate the times for median stimulus offset (r) and the median response onset (t). Results of electrocortical stimulation mapping (ESM) are represented by colored bars between pairs of electrode sites: red—involuntary tongue movement, purple—impaired spoken picture naming and auditory sentence comprehension (modified Token Test). F indicates frontal iEEG sites, and B indicates basal temporal sites. Composite of illustrations reproduced from Korzeniewska et al. (2011) with permission from Elsevier.
region sampled in at least two patients and with constant sensitivity to experimental manipulation can be considered as a region of interest (ROI). The rationale of such studies is to compensate for limited brain sampling at the individual level by including more patients. Because of the inter-individual variability of iEEG explorations, the probability of missing important nodes of the functional network under study becomes smaller as the number of patients increases. This approach has several well-known limitations though: (a) the effects reported in two different ROIs might come from two distinct groups of patients, which means that the study is in fact a collection of single-ROI studies of the kind described above (e.g. Axmacher et al., 2010b) and (b) the number of patients needed can become unrealistically high, especially when studying specific patterns of interactions: patients with electrodes over dorsolateral prefrontal cortex (DLPFC), dorsal anterior cingulate gyrus, the frontal eye fields and the intraparietal sulcus to study visual attention for instance. Once again, it is often difficult to report on more than two patients with a very specific combination of recording sites.

3.2. Individual-level or group-level statistics?

The aforementioned considerations often make group-statistics inapplicable to iEEG studies. Rather, many reports favor a detailed description of individual cases, with clear evidence that similar effects can be found in at least one other patient. Statistically, this research strategy is often combined with fixed-effects analyses, in which the effects of independent or dependent variables on HFA are tested across all trials in all subjects. In this case, intra-subject variability and inter-subject variability are treated identically, which statistically precludes a generalization beyond the investigated group of subjects (Penny and Holmes, 2004). This is in sharp contrast to non-invasive neuroimaging studies, which most often report group-statistics only, and grand-averages computed across more than 10 subjects. However, iEEG provides a rare opportunity to highlight inter-individual variability in the brain’s functional organization. Group statistics and grand-averages can be computed only in exceptional situations where the exact same brain structure is recorded in several patients. In these cases, though, it is possible to conduct random-effects analyses (with “subject” as a random variable) similar to analysis of functional MRI data (Penny and Holmes, 2004). In other words, such studies usually start with computing the average value of HFA in each subject (first level), and then conduct a parametric (or even better a non-parametric; Mari and Oostenveld, 2007) analysis across subjects (second level). In principle, this analysis strategy has the advantage that the second-level analysis tests whether the effect of an experimental manipulation or a behavioral outcome on HFA is robust against typical inter-subject differences. This allows one to generalize beyond the investigated group of subjects.

3.3. Is an epileptic human brain a good ‘model’ of the healthy human brain?

Because iEEG is always obtained in patients suffering from major brain disorders, one might rightfully wonder whether the conclusions of any intracerebral study can be extended to healthy brains. iEEG researchers have developed a series of guidelines to address this concern: (a) consider only recording sites away from the epileptic zone and signals free of epileptiform activity, such as epileptic spikes; (b) concentrate on results from functionally intact regions as assessed by neuropsychological testing and neuroimaging; (c) focus on results which can be reproduced across several patients, possibly with different seizure foci and taking different medications; (d) favor observations consistent with previous neuroimaging studies of healthy subjects.

3.4. Additional technical issues

3.4.1. A noisy environment

Experiments often have to deal with other practical issues that can potentially impair the quality of the data. One source of problems is that the clinical environment is not always designed to perform high-quality data acquisition during cognitive protocols. While most EEG researchers take great care reducing electrical noise and controlling for ambient sound and light in the recording room, clinical teams usually do not, because they do not have the same needs. Patients might hear people talking nearby or street sounds during an experiment for instance. The solution is to have a separate isolated room for cognitive experiments, as it often happens in long-standing collaborations between clinicians and researchers. But that solution requires space, and that the patient moves from one room to another during iEEG monitoring, which is not always feasible. This means that some very careful psychophysics experiments, involving stimuli at visual or auditory perceptual thresholds for instance, are particularly complicated to conduct. The clinical environment is also not always optimally designed to avoid contamination of iEEG recordings by electrical noise, and both 50 Hz and 60 Hz line-noises happen to be in the frequency range of HFA. In our experience, that noise can be extremely high in raw, nonopolar recordings. One efficient noise-canceling solution is to use a bipolar montage, that is, to use for each recording site a neighboring site as a reference. That solution is particularly well-adapted to depth-electrode recordings, where recording sites are typically separated by a few millimeters. In that case, bipolar signals are very clean and originate from the immediate vicinity of the recording site, within less than 5 mm (Lachaux et al., 2003), although one must keep in mind that neural signals recorded with the exact same amplitude on two consecutive sites cancel in such a montage. In any case, raw data must never be acquired directly using a bipolar montage, so that the experimenter can always go back to the original nonopolar signals if need be. The bipolar montage is applied offline. Bipolar montages are less adequate for 2-D cortical grids, first because the distance between sites is usually larger, which means that a bipolar montage might spread local HFA over several square-centimeters, and second because the choice of the reference for each site is less simple (most sites on the grid have eight neighbors). Laplacian transforms, a popular choice in scalp EEG research, might seem well-adapted to the geometry of cortical grids but they also spread local HFA over non-responsive sites and reduce the spatial resolution of the recordings. Alternatives include using as a reference (a) an intracranial site located in a non-responding area, for instance in the white matter, (b) a global average reference, for instance the mean of all sites with comparable size and impedance and showing no epileptiform activity, (c) a local average reference, that is, the mean of a few signals recorded in the immediate vicinity of the site of interest, (d) an external reference, such as linked mastoids. All solutions have advantages and disadvantages: for instance, an external reference might record muscular artifacts and spread them over all iEEG signals, while a global average reference might distribute over all iEEG signals HFA produced in a few sites only. The best option might be to record linked mastoids and iEEG altogether so that all options can be tested and evaluated with regard to the recording conditions.

3.4.2. Fast, but not too soon

A final constraint of iEEG research is that the timing of the experiments must be adjusted to the patients condition, in addition to the clinical schedule. It is common sense that experiments should not take place right after electrode implantation or immediately after a seizure. But how long should the experimenters wait? The collaboration with the clinical staff is
essential at that stage, because the answers really depend on the patient's recovery, which can be estimated from a judicious assessment of both the patient's neurological function and iEEG signals. Needless to say, each patient should really be considered as a generous collaborator of the research team, and be treated with great care and respect. Some patients can perform experiments the day after surgery, but most researchers choose to wait at least a few days. Our experience is that patients with depth electrodes recover faster than patients with subdural strips or even grid electrodes, probably because the latter mechanically irritate the dura much more than the former and because of the more invasive requirements of surgical implantation. Experiments can then take place for up to two weeks typically, at least with macro-electrodes. Micro-electrode recordings impose an additional time limit because microwire signals tend to degrade over time, while macro-scale iEEG signals do not.

4. Use of HFA to anticipate and prevent post-operative cognitive deficits

The objective of presurgical iEEG monitoring is not only to define the epileptogenic network as a surgical target, but also to identify cortical regions that support critical sensory, motor or cognitive functions and should be spared from surgery. The best way to anticipate post-operative deficits is to test the 'functional integrity' (Chelune, 1995) of the candidate cortical target for surgical resection. This integrity is usually evaluated by combining neuropsychological testing, functional neuroimaging, pharmacological inactivation (Wada test), and electro-cortical stimulation. None of these methods, however, have proven accurate enough to reliably anticipate the risks of surgically induced cognitive deficits. Although electro-cortical stimulation (ECS) is often considered to be the gold standard for identifying eloquent brain areas, it does not trigger interpretable manifestations in many cortical regions, especially associative regions supporting high-level cognition. For example, in one study of medial temporal lobe stimulation, 64% of the ECS evoked no subjective or overt response, even when they were strong enough to evoke an after discharge (Halgren et al., 1978a,b). It is also not clear whether ECS actually activates or disrupts cortical networks (see David et al., 2010 for a review on this topic) and strong stimulation currents carry a risk of inducing after-discharges. For these reasons, ECS – as well as the approaches cited above – may detect only the “tip of the iceberg” and may fail to reveal important functional cortical networks. Analyses of iEEG responses during cognitive tasks are emerging as a useful complement to ECS to try to address these limitations. It should be clear from the previous sections, that information processing in the neocortex generates focal, transient and task-specific HFA increases and decreases. In this section, we discuss current applications of HFA research to map cognitive functions in patients before surgery.

4.1. Fast and systematic mapping of major functional networks with HFA: the Grenoble experience

One of the authors (J.P.L.) has developed a series of short cognitive tests, lasting about 10 min each, designed to generate task-specific, and stimulus-specific HFA in several major functional systems, including the visual system, the auditory system, the motor system as well as large-scale networks subtending language, executive attention, verbal and visuo-spatial working memory (Vidal et al., 2011). The extraction pipeline of HFA has been optimized to provide a statistical mapping of HFA for each network and each recording site in a convenient visual form, only a couple of minutes after task completion. It provides an instant visual guide to the patient's functional anatomy, which is used by the clinical team to optimize the ECS procedure. For instance, a site in the auditory cortex was stimulated at 50 Hz while the patient listened to music, because it had generated music-specific HFA during the auditory localizer. During the stimulation, the patient reported a disruption of auditory perception while listening to music but not while listening to speech. Until now, the standard procedure in Grenoble during ECS of the auditory cortex comprised solely of a verbal exchange with the experimenter. Because ECS must be done at pairs of electrodes and can be quite time-consuming, it is not practical to use it to test all the potential functions that might be present at a given site. Therefore, we suggest that the fast and systematic mapping of major functional networks with HFA can be a useful and inexpensive complement to ECS, more convenient than a series of fMRI experiments for instance. In the future, we propose that HFA may potentially be used as a biomarker of healthy cortical structures to be spared from resection, provided that more studies validate that HFA actually identifies structures critical for task-performance (see Section 4.3).

4.2. BrainTV: cognitive-field mapping in the human brain

A limitation of the previous procedure, though, which also applies to ECS mapping, is that clinicians can only test functions envisioned beforehand. For instance, if a cortical site is critical for mental calculation, it might go unnoticed if the patient is never asked to use that function, either during ECS or during a standardized cognitive test. One way to avoid that pitfall is to test the reactivity of recorded sites online, as the patient is asked to engage in multiple and very diverse cognitive operations. Because of the high signal-to-noise ratio of invasive recordings, HFA induced by cognitive activity can be detected in single trials in iEEG signals (Vidal et al., 2010; Edwards et al., 2010; Miller et al., 2011). For instance, each time a patient successfully maintains a complex visual pattern in visuo-spatial working memory, an increase in HFA can be detected in the infero-temporal gyrus (Hamame et al., 2012). It is therefore possible to compute and display HFA produced by small cortical regions in real-time, and evaluate how it appears or disappears as the patient performs daily life activities. This approach has been implemented within a system called BrainTV (Lachaux et al., 2007; Lachaux, 2011): patients watch HFA from selected regions of their own brain as if they were watching a live show on TV.

Beyond its basic entertainment value, BrainTV provides a useful research tool for presurgical functional mapping. Experimenters and clinical staff can quickly formulate hypotheses about the motor, perceptual or cognitive functions supported by a given cortical site, and then test it, revising or refining it in a matter of minutes. In the example cited above, a short BrainTV session revealed that HFA in the inferior temporal gyrus was not specific to visuo-spatial working memory, but could be induced by any type of mental imagery (Hamame et al., 2012). In another example, an auditory region responsive to speech turned out to be selective to change in speaker (Lachaux et al., 2007). Examples are plenty and have often revealed unsuspected cognitive functions associated with cortical regions that were initially part of the planned resection. BrainTV can be seen as an extension of the classic receptive-field mapping procedure performed during recordings of the cat's visual cortex (Hubel and Wiesel, 1962). It is a form of 'cognitive-field mapping' during which the BrainTV experimenter searches for the cognitive process that optimally activates a given cortical region. In a second step, these initial hypotheses are followed by rigorous, controlled tests of cortical function using HFA and ECS (Lachaux et al., 2007).

In addition, BrainTV-like systems offer a unique opportunity to observe human brain dynamics online, under natural conditions: when the patient is performing informal daily activities such as...
reading the newspaper, watching TV or simply chatting with the clinical staff (Movie 1). Other approaches that rely on non-invasive measures requires that the participant stands still in a controlled environment. With BrainTV, HFA increase can be observed in the fusiform face area each time the patient looks around and finds a familiar face, and even stronger if that person starts talking.

The BrainTV set-up is an example of an application that entirely relies on the fine spatial and temporal precision of iEEG, and on the sharp association between HFA and cognition. It works, because (a) neural populations recorded by iEEG are functionally homogeneous and generate HFA each time they participate in the cognitive process at-hand and (b) because HFA can be measured and displayed in real-time (temporal precision).

Finally, BrainTV provides a path of discovery for novel brain-computer interfaces (BCI). Any efficient BCI relies on a pairing between an electrophysiological measure and a cognitive strategy that a patient can use to control that measure at will. iEEG HFA are sufficiently task-specific to serve as BCI index, as evidenced by recent studies (Leuthardt et al., 2004, 2006, 2011; Brunner et al., 2009; Hamane et al., 2012). Furthermore, the logic can be extended to the control of single-neurons: proof that human subjects can use biofeedback to selectively control brain activity even at the single neuron level has come from a recent study where subjects managed to voluntarily alter the appearance of pictures on a computer screen by controlling the activity of four selected single neurons (Cerf et al., 2010).

4.3. Participating vs. essential sites in cortical networks

One remaining question, however, is whether neural populations activated during a task – as revealed by task-induced HFA – are truly critical for task performance. What percentage of these activated populations are simply “along for the ride”, perhaps due to spreading activation of frequently associated representations? To what degree are neural representations distributed across large cortical areas or networks of areas, and how susceptible are these representations to lesions of one or more of these areas? These questions have important implications for the clinical application of iEEG studies of human cognition.

ECS provides a convenient and effective means of studying the behavioral effects of a brief, reversible, and localized cortical “lesion”, and it is widely accepted as the best available method for identifying cortical sites that are critical to function. As such, it is used to predict whether functional impairment would result if a given site were resected (Ojemann et al., 1989). However, the hyperacuity nature of the lesion makes it impossible to account for post-lesional functional reorganization and to estimate the functional reserve of other cortical areas, and there are lingering uncertainties about whether the functional lesion is limited to the site of stimulation. These concerns, as well as other practical limitations of ECS, have motivated several studies of iEEG as a potential clinical tool for preoperative functional mapping. Whereas ECS must be carried out sequentially at pairs of electrodes, iEEG can test all implanted cortical sites simultaneously. In addition, while ECS can sometimes trigger seizures that interfere with functional mapping and do not contribute to clinical goals, passive iEEG recordings do not carry this risk.

In spite of the important potential benefits of preoperative functional mapping with iEEG, the fundamental issue of participating vs. essential activation of cortical sites in iEEG, as in fMRI and PET, has inhibited its adoption into routine clinical practice pending a favorable comparison with the existing clinical gold standard. Several direct comparisons have been made between iEEG cortical mapping and ECS (Brown et al., 2008; Crone et al., 2001b; Sinai et al., 2005; Towle et al., 2008). In these comparisons, sensitivity and specificity have been used as measures of the performance of iEEG mapping as a binary classification test for the presence of cortical function, relative to the presumed gold standard of ECS. In this framework, sensitivity has been measured as the proportion of true positives (percentage of electrode sites that were both iEEG-positive and ECS-positive among all ECS-positive sites), and specificity has been measured as the proportion of true negatives (percentage of electrode sites that were both iEEG-negative and ECS-negative among all ECS-negative sites). The results of these studies have been variable, possibly due to differences in testing and analysis methods. In addition, comparisons of the two methods are not necessarily straightforward. ECS produces an all-or-none effect during language tasks, and it is difficult to know which aspect of task processing has been interrupted. iEEG produces a map of graded task-related neural responses, and a threshold for the magnitude of these responses must be chosen to compare with ECS maps.

In general, a lower threshold for activation will include more sites as iEEG-“positive” and will increase the sensitivity of iEEG with respect to ECS at the expense of a lower specificity. This tradeoff has been observed in comparisons of fMRI with ECS, particularly when maps for single tasks are compared. Such a tradeoff was observed in a comparison in 13 subjects between iEEG and ECS maps of picture naming (Sinai et al., 2005), a task that is commonly used for ECS and is one of the most susceptible to postoperative impairments. This comparison was complicated by the fact that naming could not be tested with ECS in sites responsible for spoken responses because ECS in these sites produced uncomfortable mouth sensations and movements that usually prevented task performance. Nevertheless, once this was taken into account, this study found that the specificity of iEEG mapping relative to ECS was about 78%, and that its sensitivity was 38%. The authors concluded that iEEG mapping could be used to create a preliminary map of language cortex that could later be assessed with electrocortical stimulation. In another study, Towle et al. (2008) compared iEEG responses during word repetition and recall tasks with ECS language maps and reported a sensitivity of 63%, with a specificity of 57%.

In contrast to the aforementioned studies of complex word production tasks, Sinai et al. (2009) compared iEEG maps of a speech perception task with ECS maps of speech comprehension and found a specificity of 98%, a sensitivity of 67%, and a positive predictive value (ratio of true positives to the combination of true and false positives) of 67%. The performance of iEEG mapping relative to ECS has also been excellent when mapping motor cortex (Miller et al., 2007; Brunner et al., 2009). Why have these comparisons of iEEG vs. ECS mapping been more favorable in auditory cortex and in motor cortex than in language cortex? It is possible that the cortical populations responsible for simple perceptual and motor tasks are more densely organized than are those responsible for language tasks. In particular, lexical semantic knowledge is likely represented in widely distributed cortical networks, and any given network node may contain only a portion of the overall neuronal representation. ECS may be able to acutely disrupt task performance by inhibiting such a node or by interfering with the network to which it belongs, perhaps through abnormal propagation of action potentials in passing fibers. However, the proportion of the neuronal population that is activated at this node may not be sufficient to stand out among the overall activity recorded at an iEEG electrode. This hypothesis has yet to be confirmed, however, and may require smaller, more densely spaced arrays of iEEG electrodes.

Although studies of the potential clinical utility of iEEG have primarily focused on how its maps of cortical function compare with those of ECS, it is important to acknowledge the limitations of ECS as a “gold standard”, already mentioned above, and to remember that the most important criterion to be met by any
clinical tool for functional mapping is its ability to predict and avoid post-operative neurological impairments, while maximizing the amount of ictogenic tissue that can be resected. Determining how well iEEG meets this criteria, however, will be challenging. Of course, it will require post-operative assessments of neurocognitive outcomes (relative to pre-operative assessments), as well as outcomes with respect to seizure control. A few cases of post-operative outcomes after iEEG mapping have already been reported and have suggested that it can predict post-operative language impairments (Sinai et al., 2005; Cervenka et al., 2011) as well as memory (Grunwald et al., 1998) and seizure outcome (Grunwald et al., 1999). However, because of the relatively low incidence of post-operative impairments and the number of patients that continue to have seizures in spite of surgery, rigorous tests of the accuracy of iEEG mapping will need to be done in a larger cohort of patients, perhaps requiring a multi-site clinical trial. Simple and efficient methods for iEEG mapping, such as those described above in the Grenoble experience, would greatly facilitate such a study and serve to standardize the results obtained across patients and centers. Nevertheless, the interpretation of such a study will have to take into account the fact that the entire volume of tissue resected may not be sampled comprehensively by the implanted electrodes. Likewise, the entire network of activated sites may not be resected. In spite of these limitations, however, the practical advantages of iEEG make it a compelling method that will likely be used at least for now, as a mapping tool that complements ECS.

5. Expected conceptual, methodological and technological evolution of iEEG research on cognitive HFA

As more and more research groups use iEEG to probe the fine spatio-temporal dynamics of the human brain, their scientific questions become increasingly sophisticated and require finer and broader means of investigation. iEEG itself can be used more efficiently, through faster and better analysis techniques, but it can also be improved to provide a truly multi-modal and multi-scale perspective on the large-scale neural networks supporting cognition. In this final section, we discuss several minor and major evolutions which might shape the future of iEEG cognitive research.

5.1. Beyond epilepsy: extension to deep-brain stimulations patients

In most epilepsy centers throughout the world, the number of implanted patients appears to decrease gradually across the years. This process is due to various factors, including the development of new antiepileptic drugs, the improvement of MRI hardware and signal processing algorithms (allowing to detect subtle lesions in patients which had previously been considered as non-lesional), and increased knowledge about the putative outcome of neuro-surgical interventions given specific morphological alterations. While these developments are generally beneficial for the epilepsy patients, they render the investigation of cognitive functions with iEEG increasingly difficult. On the other hand, deep brain stimulation (DBS) is currently becoming a novel therapeutic option for an increasing number of patients with various pathologies: beyond the well-established treatment of severe Parkinson’s disease by stimulation of the subthalamic nucleus (e.g. Benabid et al., 2009; Bronstein et al., 2011), this method has (among others) been recently applied to therapy-refractory patients with obsessive–compulsive disorder (Denys and Mantione, 2009; Greenberg et al., 2010), depression (Mayberg et al., 2005; Lozanno et al., 2008; Schlaeper et al., 2008), Tourette’s syndrome (Hariz and Robertson, 2010), and even Alzheimer’s disease (Laxton et al., 2010). In many cases, wires are initially externalized, which allows one to record intracranial EEG from the implanted brain regions. It is well conceivable that in the future iEEG studies on HFA may increasingly rely on such interventions in DBS patients. A disadvantage for studies of cognitive processes, however, is that electrodes are in most cases implanted only in one brain region which is likely affected by pathological processes. In some cases, electrodes may be also implanted in areas “upstream” of the main target, or in cortical structures, but the implantation is always limited to a few sites (e.g. where the depth electrodes are inserted on their way to deeper brain nuclei).

5.2. Novel analysis strategies of iEEG signals

Novel analysis strategies will be applied to iEEG data. While earlier studies of HFA focused on univariate analyses of activity in individual brain regions or on mass-univariate analyses across various regions which were each analyzed separately, there is now an increasing number of studies using bivariate measures of functional connectivity such as spectral coherence or phase synchronization, already mentioned (for a recent review, see Fell and Axmacher, 2011). In addition, some studies have started to apply effective (directional) coupling measures to iEEG, e.g. based on Granger causality or directional entropy (e.g. Gow et al., 2009). Another potential framework for analyzing iEEG data is dynamic causal modeling (DCM). Before it was adopted by Friston et al. (2003) for analysis of fMRI/BOLD signals, this framework had developed out of neural mass modeling studies designed to account for neurophysiological (EEG, MEG, EPs) phenomena based on the mean-field mathematical approach introduced by Wilson and Cowan (1973; see also Destexhe and Sejnowski, 2009a,b). DCM has more recently been applied to event-related potentials (Kiebel et al., 2006) and oscillations (Penny et al., 2009) in scalp EEG and MEG data. In the future, this method may also be applied to intracranial EEG, possibly in combination with noninvasive imaging methods (see David et al., 2008). This may be particularly interesting to constrain the neural mass models of neural activity in the investigated areas by iEEG data recorded directly within these regions.

Beyond bivariate measures, multivariate analyses allow one to address additional models of cognitive function. One emerging field of research uses pattern-classification algorithms derived from machine learning to identify the distributed pattern of category- or even stimulus-specific neural representations (Haynes and Rees, 2006). A similar technique is the analysis of representational similarity, which allows one to investigate patterns of activity even across only a few trials (Kiani et al., 2007; Kriegeskorte et al., 2008; Mormann et al., 2011). Such algorithms have already been applied to iEEG data, e.g. during memory paradigms (Manning et al., 2011). Finally, graph-theoretical measures that describe topographical properties of distributed activity patterns (Bullmore and Sporns, 2009a,b) can be applied to iEEG data.

5.3. Multimodal imaging: hold your hand out

In the near future, we expect iEEG to be combined more frequently with non-invasive neuroimaging modalities, such as fMRI or MEG, to provide a multi-scale and multi-modal understanding of task-related HFAs. There are several reasons for such developments. iEEG investigation of cognitive HFA has grown into a field of its own, which cannot continue evolving in isolation. iEEG also has its own limitations: it suffers from limited cortical sampling and applies only to patients with severe neurological or psychiatric deficits. In contrast, fMRI and MEG/iEEG provide a full coverage of healthy brains. It is therefore highly desirable to
combine the strengths of all existing techniques to provide a comprehensive view of the human brain at work.

This integration, however, requires a better understanding of how iEEG-recorded HFA is related to fMRI and MEG/EEG signals. This relationship can be investigated through the combination of invasive and non-invasive recordings. Simultaneous iEEG and MEG recordings are challenging but feasible and can provide important steps toward a multi-scale understanding of HFA. Dalal et al. (2009) have shown that HFA induced by written words in visual cortex coincide in time and anatomical origin with MEG gamma-band responses. This observation contradicted recent and influential suggestions that gamma-band responses in non-invasive recordings are mainly due to eye-movement artifacts (Yuval-Greenberg et al., 2008). In the future, we expect that more simultaneous iEEG/MEG studies will provide a platform for validating MEG source reconstruction algorithms that are capable of localizing HFA sources from noninvasive scalp-level recordings.

Simultaneous fMRI and iEEG recordings are even more challenging, but the first successful attempts have recently been published (Carmichael et al., 2010). HFA was found to be highly correlated with the BOLD signal in somatomotor cortex, in line with previous animal studies and non-simultaneous recordings in humans (Logothetis et al., 2001; Kayser et al., 2004; Niessing et al., 2005; Lachaux et al., 2007; Ojemann et al., 2010). We expect fast technological progress in the near future to make simultaneous fMRI and iEEG recordings increasingly more common.

Ideally, analysis of HFA in human intracranial EEG data should also be combined with invasive recordings in animals. Such an approach directly allows one to conduct the translational research approach described at the beginning of this article—to study a given cognitive process at various levels ranging from in vivo and in vitro electrophysiology in animals to iEEG and fMRI recordings in humans. Currently, only very few groups have adopted such a strategy, however, as it requires a genuine inter-disciplinary effort.

5.4. Multi-scale iEEG: micro-, macro-, and multimodal electrodes

We conclude this review with an important new trend in iEEG studies of human cognition: the shift toward smaller and more closely spaced iEEG electrodes for pre-surgical monitoring, including microelectrodes capable of recording single and multi-unit activity. This trend is being driven primarily by a growing appreciation among epileptologists for the importance of high frequency epileptiform activity in the localization of brain networks responsible for medically refractory seizures. Until now intracranial EEG recordings and stimulation in common clinical practice have employed relatively large electrodes, i.e. “macro-electrodes”, with surface areas ranging from 1.25 to 4 mm², typically configured in linear or rectangular arrays with inter-electrode spacing (center-to-center) ranging from 2.2 (depth) to 10 mm (subdural). Until recently the sizes and configurations of these electrodes have been constrained in large part by the technological limitations of amplifiers commonly used in long-term epilepsy monitoring units, i.e. low input impedances, sampling rates, and channel counts. Advances in recording technology, however, are rapidly removing these limitations and making it possible to record from the same cortical territory with denser arrays capable of recording human neuronal populations at increasingly finer detail.

Although technological barriers to high-density iEEG are quickly disappearing, most clinicians to date have been satisfied with traditional electrode arrays and recording technology to achieve the primary clinical goal of localizing the epileptogenic, as well as eloquent, cortex. Studies at several epilepsy centers have recently suggested, however, that smaller, more closely spaced electrodes may offer benefits to patients beyond those offered by traditional macroelectrodes, and commercial medical vendors (e.g. Adtech, PMT) have begun to offer FDA-approved hybrid (macro- micro) electrode arrays in which sub-arrays (1-mm spacing) of “microelectrodes” (e.g. 0.561 mm diameter, surface area 0.25 mm²) are interposed between traditional macro-electrodes. Most importantly, both animal and human studies have indicated that high frequency oscillations recorded from microelectrodes are reliable indices of the epileptogenic zone and may improve localization of the seizure focus and thereby the outcome of epilepsy surgery. Indeed, the rapidly growing number of studies to this effect are comprehensively represented in the other contributions to this special issue. These studies have occurred in parallel with and largely independently of the studies using HFA to study normal human cortical function.

The question of what is the optimum electrode size to record epileptiform HFOs remains open. In the first systematic study to date of epileptiform high frequency oscillations (HFOs) recorded at different spatial scales, Worrell et al. (2008) performed combined recordings with standard depth electrodes (contact area 9.4 mm²) and with microelectrodes (40 μm diameter) placed 1–5 mm away. They concluded that both electrode types could record epileptiform HFOs, but that high frequencies were better recorded with the microelectrodes. A subsequent study using the same depth electrodes as well as micro- and micro-contact electrodes in subdural electrodes found no excess high frequency activity in the subdural microelectrodes (Blanco et al., 2011), possibly because they were too far away from the small neuronal ensembles generating this activity. Therefore, very high-frequency activity might be picked up only by micro-electrode contacts on depth electrodes penetrating the parenchyma. A series of studies from the Montreal Neurological Institute using larger electrodes with a diameter of ~1 mm and a surface area of 0.8 mm², still smaller than most commonly used electrodes, have also recorded epileptiform HFOs during (Jirsch et al., 2006) and between seizures (Crepon et al., 2010; Jacobs et al., 2008; Urestarazu et al., 2007), and have demonstrated a significant correlation between the resection of regions with epileptiform HFOs and postsurgical seizure-free outcomes (Jacobs et al., 2010; Ochi et al., 2007). In most cases, epileptiform HFOs appeared to be a better index of the epileptogenic zone than traditional epileptiform discharges seen in lower frequencies.

To date, the sensitivity and specificity of HFA-based ECoG mapping relative to electro-cortical stimulation (ECS) has been excellent in sensorimotor cortex (Brunner et al., 2009; Miller et al., 2009a,b,c) and in auditory cortex (Sinai et al., 2009), where functional anatomy is densely and predictably organized. In language cortex, however, where function is more widely distributed, the sensitivity of ECoG with respect to ECS has been less than optimal (Sinai et al., 2005; Towle et al., 2008). It is possible that higher density iEEG recordings such as those afforded by newly available hybrid electrodes with both macro- and micro-electrode elements, will improve the sensitivity of ECoG with respect to ECS. There is growing evidence for such a hypothesis from studies using subdural electrode arrays with 4 mm spacing instead of the 1 cm spacing that has been more commonly used in epidural grids. In recordings of high gamma responses to speech stimuli in human auditory association cortex, several studies have found evidence for functional specialization that is sufficient to discriminate cortical regions responsible for different phonetic category boundaries (Chang et al., 2010), as well as cortical regions responding to self-generated vs. externally generated speech (Flinker et al., 2010, 2011).

Because micro-electrodes emphasize the activity of neuronal populations in their immediate vicinity, cognitive neuroscientists may be able to use dense arrays of these electrodes to identify neuronal populations with even more specific functional
responsibilities than have been observed to date. This would be particularly useful when attempting to decode the information represented in these populations as in, for example, brain-machine interfaces driving brain-controlled prostheses. Furthermore, it may be possible to determine the density of these populations in any given patch of cortex and perhaps better estimate the likely functional consequences of its resection.

Beyond their considerations for functional mapping and brain–computer interfaces, iEEG recordings combining micro- and macro-electrodes may also provide a vital opportunity to better understand the basic physiological mechanisms of HFA and other macroscopic EEG phenomena that have been used for human cognitive neuroscience. For example, it may be possible to better test whether neuronal populations with different functional selectivities can be discriminated by the frequency of their gamma responses, as described above and recently supported by the study of Gaona et al. (2011). Furthermore, cognitive neuroscientists inspired by the success of human single-unit studies (e.g. Engel et al., 2005; Quiroga et al., 2008), will undoubtedly want to investigate the relationships between single and multi-unit activities (firing rates and spike timing) and the aggregate local field potentials of nearby neuronal populations. Recent evidence suggests that extracellular HFA is not only an emergent result of neural population activity, but may also exert a causal role on neural activity itself (Frohlich and McCormick, 2010; Anastassiou et al., 2011). This question could be ideally addressed with a combination of macro- and microelectrode recordings.

The field of iEEG will also undoubtedly be boosted by a new generation of intracranial electrodes, both multi-scale and multimodal, such as the Neuropore (Grand et al., 2010) or the optode (Keller et al., 2009). This new generation of iEEG recordings is expected to provide a firm experimental basis for understanding the links between HFA, local synchronization processes, and metabolic activation. Yet another exciting possibility is the combination of iEEG recordings with voltammetry or microdialysis (e.g. Fried et al., 2001). This approach allows one to test predictions about the effect of specific neurotransmitters on HFA in a given brain region, e.g. the effect of dopamine release on hippocampal activity (Lisman and Grace, 2005; Axmacher et al., 2010b; Lisman et al., 2011), but it also raises very complex safety issues that should be resolved before use.

Acknowledgments

J.P.L. was supported by two grants from the French National Research Agency (ANR: OpenVibe2 and MLA). NA was supported by an Emmy Noether grant from the Deutsche Forschungsgemeinschaft. F.M. was supported by grant MO930/4-1 from the DFG. N.E.C. was supported by NINDS R01-40596. EH was supported by the US NIH (NS18741) and by INSERM. We thank all the patients who participated in the studies cited in this review. N.E.C. thanks Anna Korzeniewska, Ph.D. for Fig. 6. J.P.L. thanks Tomas Ossandón for Fig. 5.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.pneurobio.2012.06.008.

References


Ray, S., Crooke, N.E., Niebur, E., Franaszczuk, P.J., Hsiao, S.S., 2008. Neural correlates of high-gamma oscillations (60–200 Hz) in macaque local field potentials and