MEG and Language
Suzanne Dikker, M Florencia Assaneo, Laura Gwilliams, Lin Wang, Anne Kösem

To cite this version:
Suzanne Dikker, M Florencia Assaneo, Laura Gwilliams, Lin Wang, Anne Kösem. MEG and Language. 2019. hal-02265485

HAL Id: hal-02265485
https://hal.archives-ouvertes.fr/hal-02265485
Submitted on 15 Aug 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
MEG and Language:

Using Magnetoencephalography to Study the Neural Basis of Language

Suzanne Dikker ¹, Florencia Assaneo ¹, Laura Gwilliams ¹, Lin Wang ² & Anne Kösem ³

(1) Department of Psychology, New York University, USA
(2) Department of Psychiatry and the Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, USA
(3) Postdoctoral researcher, Lyon Neuroscience Research Center (CRNL), Lyon, France

This text is adapted from MEG and Language by the same authors, submitted to an issue of Neuroimaging Clinics on Magnetoencephalography, edited by Drs. Roland Lee and Mingxiong Huang

All authors contributed equally to this work

Synopsis

We provide an introductory overview of research that uses magnetoencephalography (MEG) to understand the brain basis of human language. The cognitive processes and brain networks that have been implicated in written and spoken language comprehension and production are discussed in relation to different methodologies: we briefly review event-related brain responses, research on the coupling of neural oscillations to speech, oscillatory coupling between brain regions (e.g., auditory-motor coupling), and neural decoding approaches in naturalistic language comprehension. We end with a short section on the clinical relevance of MEG language research, focusing on dyslexia and specific language impairment.
Introduction

Language is among the most complex of human cognitive systems, yet its processing is extremely automated and fast: Both behavioral and neurophysiological studies suggest that within 600 milliseconds of a word’s onset its sensory properties have been analyzed, its grammatical and semantic features have been retrieved from memory, and it has been integrated into the ongoing discourse. In this chapter, we provide a brief overview, summarized in Figure 1, of research that has capitalized on MEG’s spatio-temporal resolution to capture the neural dynamics that support linguistic operations. First, we briefly review event-related designs using single words or simple, well-controlled phrases, which have helped characterize the building blocks of language and its processing stages, from identifying individual sounds or letters to sentence-level grammar. We then discuss more recent methodological approaches exploring neural oscillations and decoding techniques, which can be applied to more naturalistic language. These tools have allowed researchers to study how linguistically relevant brain responses are coupled between brain regions and across modalities, and to capture language-relevant brain responses that are highly distributed across space and vary in timing. We conclude with a brief section exemplifying the clinical relevance of the reviewed research.

Evoked responses to linguistic input

Early brain responses to written words

In a now classic study, Tarkiainen and colleagues \(^1\) compared brain responses to symbols and letter strings and identified two early event-related MEG responses to visual word recognition \(^1\). Later studies have built on this work to identify early language-related MEG components, or event-related fields (ERFs) \(^2,3\). As listed in Figure 1, the so-called visual M100 response has been associated with low-level visual feature processing \(^4\), the M130 component shows sensitivity to orthographic features \(^3\) (e.g., how often two letters co-occur in written words), and the M170 response is sensitive to the morphological properties of words \(^5\). For example, bimorphemic words elicit a higher M170 amplitude than orthographically matched monomorphemic words (“farm-er” has two morphemes, or meaningful units, vs “corner”, which consists of one morpheme but also
ends in “-er”). fMRI evidence from manipulations similar to those affecting the M170 suggests that the response is generated in left and right fusiform gyri.

Identifying phonetic and morphemic information in spoken language

Systematic early responses to spoken language can also be identified. Relative to the onset of each speech sound, evoked responses at 50 ms in primary auditory cortex are modulated by low-level spectro-temporal properties of the input, followed by a 100 ms response in superior temporal gyrus. This latter response has been associated with the mapping of variable acoustic information onto more stable phonetic features (e.g., the acoustics of [p] vary between speakers and contexts, like in [plant] vs. [park], but these different sounds are mapped onto a single meaningful phoneme in the perceiver’s brain).

Having identified the phonemes, these are then mapped onto words (via morphemes, see above). This process has been explained through the prevalent cohort model of spoken word recognition (Figure 2): as a listener hears each phoneme of a word (i.e., the spoken equivalent of a letter), this information is used to narrow down the cohort of words that are consistent with the input. When just one word remains, this is the winner of the lexical competition, and thus the word that is recognized. Previous MEG studies have found evidence in favor of this mechanism by
tracking responses to each speech sound in the word, and correlating activity at around 200 ms after phoneme onset in superior temporal gyrus with the number of remaining morphological candidates \(^{10,12,13}\) (Figure 2).

**Figure 2. Schematic of spoken word recognition.** As the word unfolds, with each phoneme (gray bars), possible lexical candidates for recognition are ruled out. This process continues until only one word remains consistent with the input (e.g., “parakeet”). Notice that the word can be uniquely identified before word offset -- at the “k” phoneme. Activity in left superior temporal gyrus has been shown to track the probability of each phoneme in a word (“phoneme surprisal”), as well as the relative probability of remaining lexical candidates (“cohort entropy”). Figure courtesy Laura Gwilliams.

Deriving the meaning of words in isolation and in context

Once the words have been identified, their meaning can be derived. A large body of MEG studies have compared evoked brain responses to words differing in lexical properties (words vs. non-words; abstract vs. concrete words \(^2\)), preceding word contexts (e.g., “dog” preceded by “cat” vs. “dog” preceded by “table” \(^{14,15}\)), or sentence contexts (e.g., “mountain” vs. “tulip” in “The climbers finally reached the top of the ...” \(^{16–18}\); Figure 3A). When semantic features of a word are easier to access, smaller evoked MEG responses are observed approximately 300-500 ms after word onset (corresponding to the N400 in EEG \(^{19}\)). This N400m effect has been consistently localized within the left lateral temporal cortex, including superior and middle temporal regions (Figure 3B).
Figure 3. Evoked responses to congruent vs. incongruent words in sentence contexts. (A) The planar gradient of the ERFs time-locked to the onset of the critical words at a representative left temporal sensor. Approximately between 200ms and 700ms after the critical words, the incongruent words (e.g. “tulip”) elicited larger amplitudes than the congruent words (e.g. “mountain”). (B) Source localization of the N400m effect using MNE. Figure adapted, with permission, from Wang et al. (2012) Beta oscillations relate to the N400m during language comprehension. Human Brain Mapping, 33(12), 2898–2912.

This is in line with the well-established role of these regions in representing lexico-semantic information, as revealed by fMRI. Importantly, while early ERFs are modality specific, semantic activation appears to be modality-independent: overlapping brain regions are activated during the N400m time window for both written and spoken language. In addition to the left temporal cortex, a number of MEG studies report increased activation within the left inferior frontal cortex for semantically incongruous words. This region has been suggested to play a role in selecting and controlling lexical retrieval under contextual influence and in unifying multiple sources of information.

Combining words: syntax and sentence-level semantics

While the bulk of studies focus on processing word meaning, some research also examines sentence-level combinatorics. For example, when comparing the word “boat” in either a word list or as part of the phrase “red boat”, the latter leads to linguistic composition-related activity in the left anterior temporal lobe and ventromedial prefrontal cortex (e.g., see Figure 1). And “going” vs. “go” in a sentence like “She will going/go to the bakery for a loaf of bread.” elicit a stronger response after 600 ms, consistently source-localized to temporal regions, corresponding to the P600 effect observed in EEG research.

Syntactically incongruous words sometimes also elicit stronger early sensory responses than grammatical sentences, as early as 100 ms post-word onset. Recall that research investigating words in isolation yielded no evidence that the brain is sensitive to lexical or
grammatical properties in sensory cortices. This discrepancy between word-level and sentence-level findings has been explained in terms of a prediction-based account: words and grammar can be anticipated (and thus preactivated) based on the preceding context, allowing for a more efficient detection of linguistically relevant features. For example, the brain may generate estimates of the likely physical appearance of upcoming words based on grammar-based predictions, and words that do not ‘look’ or ‘sound’ like the expected grammatical category then show increased early sensory responses. Many of the studies cited above in relation to lexical-semantic processing have similarly accounted for N400m effects in terms of prediction\textsuperscript{14,33,34}.

Beyond ERFs and toward naturalistic experimentation

In recent years, a shift has been made towards a more naturalistic experimental setup (e.g.,\textsuperscript{35}) where participants listen to continuous speech, such as stories, rather than being presented with isolated words or short sentences. As the evoked response to natural language at a certain time point may reflect a cascade of processes initiated at different (and overlapping) moments in time, researchers have resorted to other analysis techniques to study continuous speech.

The role of neural oscillations in language

Tracking continuous speech: brain-to-stimulus synchrony

There is strong evidence that brain oscillations synchronize to the temporal regularities of speech during listening\textsuperscript{36–39}: low-frequency neural oscillations in the theta and delta range (1-8 Hz) synchronize to the dynamics of the speech envelope associated with syllabic and phrasal rates presentation respectively, and high-frequency neural activity in the gamma range (30-40 Hz) follow the fine-grained temporal dynamics related to phonetic features\textsuperscript{40} (Figure 4A).

Neural oscillations have been experimentally linked to the understanding of speech (i.e., beyond mere speech acoustics): brain-to-speech synchronization is stronger when speech is intelligible (Figure 4B)\textsuperscript{41–43}, and brain oscillations in the delta and gamma range synchronize to the lexical and grammatical structures of spoken sentences\textsuperscript{44–46}. In multi-talker settings, neural oscillations synchronize to the dynamics of the attended speaker (the so-called ‘cocktail party
effect’). Importantly, the strength of the synchronization indicates how successful the auditory scene analysis is performed \(47-49\) (Figure 4C).

Neural oscillations are hypothesized to play a crucial role during speech parsing by defining the temporal boundaries between linguistic items within the continuous acoustic signal \(40\). Importantly, brain-to-speech synchrony is suggested to be a predictive mechanism, parsing and structuring events from continuous speech by building temporal expectations on the upcoming auditory input \(50,51\). In support of this view, neural oscillations entrain (synchronize) to the syllabic rate of speech in a sustained manner, and neural entrainment to ongoing speech is dependent on the rate of preceding speech. Importantly, brain-to-speech synchrony at specific rates directly affects how words are understood \(52\) (Figure 4D).

![Figure 4. Brain-to stimulus synchrony and speech processing. (A) Brain signal synchronizes to the dynamics of speech. (B) Brain-to-stimulus synchrony is stronger when speech is intelligible\(^{41-43}\). (C) Brain oscillations follow attended speech in multi-talker settings\(^{47-49}\). (D) Brain oscillations are influenced by contextual speech rhythm information\(^{52}\). Figure courtesy Anne Kösem.](image)

The role of neural oscillations in deriving meaning

Examining oscillatory activity has also allowed researchers to closely track the cortical dynamics underlying semantic processing in event-related designs, substantially enriching the information obtained by studying ERFs. For example, during sentence comprehension, desynchronization of neural oscillations in the alpha (8-12 Hz) and beta (12-20 Hz) range within the left temporal and
Frontal regions are observed when processing semantically incongruent (versus congruent) words in sentence contexts \(^{18,28,53}\) or for words appearing in sentences vs. word lists \(^{54}\). The desynchronized alpha/beta activity might reflect the engagement of task-relevant brain regions to support sentence-level processing. Furthermore, connectivity analyses in MEG studies have revealed functional connectivity patterns between the left inferior frontal and temporal cortex during sentence comprehension. Granger causality analysis suggests that alpha activity supports information transfer from temporal to frontal regions, whereas beta activity supports information transfer from frontal to temporal regions \(^{55}\). The connectivity between the left frontal and temporal regions was also found by synchronized beta and low-gamma oscillations for processing unexpected compared to expected sentence-final words \(^{56}\). Moreover, cross-frequency connectivity between gamma power within the left prefrontal region and alpha power within the left temporal region was reported for processing expected but not unexpected sentence-final words, both before and after the words were presented \(^{53}\). These findings suggest that the communication between the left inferior frontal and temporal cortex is supported by different patterns of oscillatory activity. Further studies are needed to systematically examine how the communication between different brain regions (e.g., feedforward vs. feedback control) is realized via different patterns of synchronization (e.g., phase-locking, amplitude synchronization, phase-amplitude coupling) in different frequency bands \(^{53,57}\).

**Auditory-motor interaction in speech**

In the middle of the past century, researchers observed that phonemes cannot be unequivocally defined in the acoustic space, and proposed that phoneme perception instead occurs in motor space: incoming speech sounds are mapped to invariant neuromotor commands \(^{58,59}\). This hypothesis, dubbed the *motor theory of speech* perception, has been strongly criticized \(^{60–65}\), but less stringent versions are supported by findings that passive listening to speech activates areas involved in speech production \(^{66,67}\). MEG data show that oscillations generated in left inferior frontal and precentral gyri (areas typically involved in speech production) modulate the phase of low-frequency activity in left auditory regions significantly stronger when speech is intelligible than when it is unintelligible (e.g. backward speech). This top-down control leads to a better tracking of the speech envelope by the auditory cortex activity \(^{68,69}\), suggesting that motor areas help enhance auditory temporal prediction during speech processing \(^{50}\).
In recent MEG research where participants passively listened to rhythmic strings of syllables, the synchronization between auditory and motor regions was found to be highest when syllables were presented at a rate of 4.5 syllables per second (Figure 5), corresponding to the mean syllable rate of natural speech across languages. One plausible explanation is that speech production regions behave as a neural oscillator—a system capable of generating oscillations at its own natural frequency and showing entrainment to a rhythmic stimulus only if the external frequency is close to its natural one—with a natural frequency close to 4.5 Hz. Thus, the temporal patterns of speech could emerge as a consequence of the intrinsic rhythms of cortical areas.

Figure 5. Auditory-motor synchronization during speech perception. (A) Experimental data. Motor (red) and auditory (blue) cortical activity were recorded while participants passively listened to rhythmic train of syllables at different rates. The synchronization between cortices significantly increased from baseline just if syllables were presented at 2.5, 3.5 and 4.5 syllables/sec and was enhanced at the central condition (4.5 syllable/sec). (B) Model output. Motor cortex (red box) was modeled through a set of Wilson-Cowan equations representing an inhibitory-excitatory network, and the excitatory population receives the auditory cortex activity (blue signal) as input.

The Wilson-Cowan is a biophysically inspired model, which behaves as a neural oscillator. Numerical simulation obtained by setting the natural frequency of the oscillator at 4.5 Hz reproduced the experimental pattern of auditory-motor synchrony. Figure adapted, with permission, from Assaneo MF, Poeppel D. The coupling between auditory and motor cortices is rate-restricted: Evidence for an intrinsic speech-motor rhythm. Science Advances. 2018;4(2).

Decoding approaches in language research

Another approach that has gained popularity in recent years is “multivariate” or “decoding” analyses. For MEG, this typically involves using the activity pattern across all sensors (not just one location at a time; hence, “multi-” variate), in order to read out the stimulus property as it is
encoded in the neural response (hence, “decoding”). The algorithms for this approach have been largely borrowed from the machine-learning community, and can provide an increased sensitivity in situations where the neural processes do not necessarily evoke a strong focal amplitude modulation. For example, evidence from a decoding analysis on EEG-MEG data aimed at differentiating responses to words of different semantic categories (e.g. living vs. non-living), suggests that relevant responses are not only highly distributed across space, but also vary in timing. Further, a recent study demonstrated the ability to decode the identity of a word, before it begins, when that word is highly predictable in context. This was achieved by examining the spatial pattern of brain activity prior to the presentation of predicted words. Overall, these studies show that applying multivariate analysis to the MEG data paves the way for studying language processing in cases where classic evoked-responses analysis would not be sensitive to the true underlying processing differences.

Clinical relevance: dyslexia and specific language impairment

Disrupted evoked responses to speech and text has been associated with several language deficits. For example, dyslexic participants exhibit atypical early evoked responses to letters, and N400m responses measured in infancy can predict reading speed in adolescence for children at risk for dyslexia. Disruption of early neural processing to speech sounds has also been linked to dyslexia, as well as specific language impairment (SLI).

The study of neural oscillations and their link to speech and language disorders is also receiving increasing interest. Hearing impairment affects alpha oscillations and speech-brain synchronization; both oscillatory deficits are associated with the degree of hearing loss in noisy environments. Dyslexia and SLI have been proposed to originate from abnormal neural oscillatory profiles in response to speech. Children who have inefficient brain-to-speech synchronization may have a harder time segmenting the acoustic signal, leading to language difficulties. Dyslexics show stronger synchronization to fast auditory rhythms in the gamma range (40 Hz) as compared to controls, suggesting disrupted auditory timing processing for phoneme perception. Brain synchronization to the slower dynamics of the speech envelope is also disrupted in dyslexia and SLI.
To remediate language deficits, rhythm training programs have been tested to improve speech timing perception based on finger-tapping or musical exposure\(^{87,88}\). Finger-tapping and musical exposure have both been associated with increased recruitment of neural oscillations\(^{89,90}\). Ongoing research is currently investigating the direct link between neural oscillatory activity and speech perception improvement during training.

References


30. Dikker S, Rabagliati H, Farmer TA, Pylkkänen L. Early occipital sensitivity to syntactic category is


65. Skipper JI, Devlin JT, Lametti DR. The hearing ear is always found close to the speaking tongue: Review of the role of the motor system in speech perception. *Brain Lang*. 2017;164:77-105.


