
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Author(s): Hultén, Annika & Karvonen, Leena & Laine, Matti & Salmelin, Riitta
Title: Producing Speech with a Newly Learned Morphosyntax and Vocabulary: An Magnetoencephalography Study
Year: 2014
Version: Final published version

Please cite the original version:

Hultén, Annika & Karvonen, Leena & Laine, Matti & Salmelin, Riitta. 2014. Producing Speech with a Newly Learned Morphosyntax and Vocabulary: An Magnetoencephalography Study. *Journal of Cognitive Neuroscience*. Volume 26, Issue 8. P. 1721-1735. 0898-929X (printed). DOI: 10.1162/jocn_a_00558.

Rights: © 2014 MIT Press. This is the final version of the article: Hultén, Annika ; Karvonen, Leena ; Laine, Matti ; Salmelin, Riitta. 2014. Producing Speech with a Newly Learned Morphosyntax and Vocabulary: An Magnetoencephalography Study. *Journal of Cognitive Neuroscience*. Volume 26, Issue 8. P. 1721-1735. ISSN 0898-929X (printed). DOI: 10.1162/jocn_a_00558, which has been published in final form at http://www.mitpressjournals.org/doi/abs/10.1162/jocn_a_00558#.VVlqF2OkuSo

All material supplied via Aaltodoc is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

Producing Speech with a Newly Learned Morphosyntax and Vocabulary: An Magnetoencephalography Study

Annika Hultén^{1,2}, Leena Karvonen¹, Matti Laine², and Riitta Salmelin¹

Abstract

■ Ten participants learned a miniature language (Anigram), which they later employed to verbally describe a pictured event. Using magnetoencephalography, the cortical dynamics of sentence production in Anigram was compared with that in the native tongue from the preparation phase up to the production of the final word. At the preparation phase, a cartoon image with two animals prompted the participants to plan either the corresponding simple sentence (e.g., “the bear hits the lion”) or a grammar-free list of the two nouns (“the bear, the lion”). For the newly learned language, this stage induced stronger left angular and adjacent inferior parietal activations than for the native language, likely reflecting a higher load on lexical retrieval and STM storage. The preparation phase was followed by a cloze

task where the participants were prompted to produce the last word of the sentence or word sequence. Production of the sentence-final word required retrieval of rule-based inflectional morphology and was accompanied by increased activation of the left middle superior temporal cortex that did not differ between the two languages. Activation of the right temporal cortex during the cloze task suggested that this area plays a role in integrating word meanings into the sentence frame. The present results indicate that, after just a few days of exposure, the newly learned language harnesses the neural resources for multiword production much the same way as the native tongue and that the left and right temporal cortices seem to have functionally different roles in this processing. ■

INTRODUCTION

An essential goal in learning a new language is to be able to use it for carrying a conversation. Yet, learning to verbally express yourself in a foreign language generally takes longer than the time required for good comprehension skills (Clark & Hecht, 1983). Real-life conversation involves a multitude of cognitive processes but at its core lies sentence-level production that requires command of both the vocabulary and syntactic/morphosyntactic operations. Nevertheless, our knowledge on the neural substrates of language production beyond single-word level is still quite limited, as they have been addressed only in a handful of studies (Menenti, Petersson, & Hagoort, 2012; Menenti, Gierhan, Segaert, & Hagoort, 2011; Brownsett & Wise, 2010; Dhanjal, Handunnetthi, Patel, & Wise, 2008; Awad, Warren, Scott, Turkheimer, & Wise, 2007; Golestani et al., 2006; Haller, Radue, Erb, Grodd, & Kircher, 2005; Blank, Scott, Murphy, Warburton, & Wise, 2002; Indefrey et al., 2001; Kircher, Brammer, Williams, & McGuire, 2000). Even less is known about the neural underpinnings of language production in native versus foreign language (Golestani et al., 2006; Kim, Relkin, Lee, & Hirsch, 1997). This study explores the cortical dynamics of second language production at sentence level using magnetoencephalography (MEG).

There is experimental evidence that retrieval of newly learned words in the native language shares neural correlates with retrieval of well-known familiar words, and the newly learned words can be successfully integrated into the adult mental lexicon for long-term permanent storage (Hultén, Laaksonen, Vihla, Laine, & Salmelin, 2010; Davis, Di Betta, Macdonald, & Gaskell, 2009; Hultén, Vihla, Laine, & Salmelin, 2009; Dobel, Gumnior, Bölte, & Zwitserlood, 2007). As regards acquisition of an artificial grammar, receptive tasks have suggested that the novel grammar is also processed in a native-like manner (Morgan-Short, Finger, Grey, & Ullman, 2012; Morgan-Short, Steinhauer, Sanz, & Ullman, 2011; Friederici, Steinhauer, & Pfeifer, 2002). However, it is possible that when a novel vocabulary and a novel grammar are actively combined in language production, the two may interact, rendering production not only more taxing but also qualitatively different. According to the declarative/procedural model of lexicon and grammar (Ullman, 2004), adult language learners process both grammar and vocabulary using their declarative memory system, whereas native speakers recruit procedural memory for syntactical processing and only engage the declarative memory for vocabulary storage (Ullman, 2001). However, there is mounting evidence that the first and second languages as a whole share the same neural substrates (Perani & Abutalebi, 2005), and it has been suggested that any differences between the two are more likely to be because of factors such as proficiency level or transfer between

¹Aalto University, ²Åbo Akademi University

the languages (Kotz, 2009; Rodríguez-Fornells, Cunillera, Mestres-Missé, & de Diego-Balaguer, 2009; Friederici et al., 2002). Against this background, it seems likely that a novel vocabulary and novel grammatical operations are processed in the same general brain network that is involved in native language use. However, differences between early and late language learning may yet become apparent in the more demanding domain of sentence-level language production.

The neural correlates of narrative speech production are technically difficult to assess (Price, 2010), and they have been investigated in only a few studies, using quite diverse paradigms. Therefore, there is no clear understanding of the spatiotemporal dynamics of narrative speech production even in the native language (control) condition. Hemodynamic studies focusing on the left inferior frontal gyrus have assigned this area a role in syntactic processing also during production, as it appears to be more active when generating sentences from a given set of words than when simply reading them (Golestani et al., 2006; Haller et al., 2005) or when describing moving geometrical shapes using sentences rather than noun phrases or single words (Indefrey et al., 2001). These findings are in line with the traditional neuropsychological view that has linked left frontal lesions with agrammatic speech production in aphasia (e.g., Goodglass & Kaplan, 1983). In whole-brain analyses, production of a written or spoken narrative has been shown to activate frontal and temporal areas as well as the TPJ (including the left angular gyrus), when contrasting propositional and nonpropositional output (e.g., counting vs. syllable repetition; Brownsett & Wise, 2010; Awad et al., 2007; Blank et al., 2002) or when identifying activation correlated to speech rate (Kircher et al., 2000). At the single-word level, the literature is more extensive, including studies that have used electrophysiological measures to provide the high temporal resolution needed for capturing the rapidly unfolding neural events in language production. For example, the sequence of activation elicited by naming a pictured object is relatively well known, proceeding from the occipital cortex (<200 msec) to the inferior parietal and posterior temporal cortices (>200) and lastly to the frontal cortex (Vihla, Laine, & Salmelin, 2006; Levelt, Praamstra, Meyer, Helenius, & Salmelin, 1998; Salmelin, Hari, Lounasmaa, & Sams, 1994).

In sentence-level processing, an added sentential meaning arises from the combination of words and syntax that together convey a proposition. Different grammars may have different means of expressing these combinatorial links. For example, where English uses mainly word order, Finnish relies heavily on inflectional endings to signal the same thematic roles. Meaningful word combination in sentence production in English has been linked to the right lateral temporal cortex. Kircher, Brammer, Tous Andreu, Williams, and McGuire (2001) found increased activation in that area in a cloze task where free generation of a semantically most fitting sentence-final word was contrasted to overt reading or selection of the sentence-final target word among two alternatives. In an MEG study

where participants matched minimal linguistic phrases (“red boat”) to an upcoming picture, activity in the left angular gyrus and anterior temporal lobe was associated with combinatorial processing (Bemis & Pylkkanen, 2013). In Dutch, using fMRI, mapping pictorial reference to the semantic structure of a sentence was shown to activate a widespread network including the left middle frontal gyri and the bilateral superior parietal and posterior temporal gyri (Menenti et al., 2012). However, it remains unclear whether similar processes underlie combinatorial operations in languages that implement semantic relations primarily through morphosyntactic means instead of word order.

The present study seeks an answer to two questions that relate to the neural mechanisms underlying second language processing when thematic roles are expressed by morphosyntax. First, if the morphosyntactic structure of a newly learned language deviates from that of the native tongue, do the underlying neural substrates differ functionally from those of the native tongue? Second, which cortical systems are engaged when we prepare an utterance and retrieve lexical information with versus without sentence context and corresponding morphosyntactic structure?

By using an artificial language instead of a real language, we gain full control of the amount of exposure to the language and its linguistic characteristics. We seek to minimize transfer effects by creating a miniature language where neither the grammar nor the vocabulary resembles that of the participants’ native tongue, Finnish. Furthermore, participants are trained to a high level of proficiency in the novel language, which is necessary for fluent output in the task. By comparing production of the native language versus the novel language, the participants serve as their own controls for early versus late language acquisition. A behavioral follow-up test 6 months later will additionally inform on the stability of the relatively short but intense learning period and the long-term maintenance of syntactical and word-level knowledge.

The artificial language is trained in a procedure that emulates formal classroom learning and gives explicit information of the grammatical rules and vocabulary. This type of learning is ecologically valid for adults, and it also ensures that all participants start off with the same information, independent of their ability to infer explicit rules from the input. Recent findings suggest that native-like neural processing can be achieved both through explicit classroom-like training and with implicit immersion-like training (Morgan-Short et al., 2012).

We use a production task to tap into proficient language usage, but the present task is also a continuation to the studies on single-word learning using picture naming (Hultén et al., 2009, 2010; Grönholm, Rinne, Vorobyev, & Laine, 2005; Raboyeau et al., 2004). In natural language usage, single words need to be expressed as part of a coherent sentence within the current discourse, but sentence production remains difficult to study at the neural

level. Thus, although the production task is motivated from the perspective of language learning, some theoretical and methodological aspects of the sentence-level task are relevant also beyond the language learning domain.

At a first glance, the best way to study the neural underpinnings of sentence production would be to register brain activation during continuous speech production (for this procedure in fMRI, see, e.g., Kircher et al., 2000). However, such an approach makes it difficult to assess which underlying cognitive processes are involved (Haller et al., 2005). It also leads to practical problems in electrophysiological measures with regard to baseline and measurement artifacts from the moving tongue and face muscles. As theories on speech production distinguish between preparation and execution of speech output (Levelt, Roelofs, & Meyer, 1999; Dell, 1986), using a task setup that mimics this major division seems motivated. A preparation phase is free from output-related artifacts, and these artifacts can be minimized also for the execution phase by employing delayed production. We use pictured events to prompt controlled processing of multiword utterances that either have a morphosyntactic structure (sentence) or not (separate nouns). A fixed task sequence is used to separate initial preparation for output from a subsequent cloze task. In the Preparation phase, participants are asked to think of the sentence or word pair that corresponds to the presented cartoon. This is followed by the Cloze task which calls for covert production of the final word. To ensure task compliance, the last word is then produced overtly after a short delay. This design allows for time-locked evoked responses to be recorded to both the Preparation phase and to each word of the phrase in the Cloze task. The setup is used here to probe cortical activation patterns in both native language and novel language utterances. If grammatical processing in the first and second language relies on different memory systems, we could expect differences in neural representation of the two languages (in contrast to word processing where the first and second language should rely on the same memory system). If such patterns are present in language production, they might be more readily tractable with the time-sensitive MEG than with the relatively slow temporal dynamics of hemodynamic measures. Furthermore, as our study contrasts production of morphosyntactically organized sentences versus noun sequences, it may shed light on the neural correlates of sentence-level combinatorial processes and production of morphosyntax.

METHODS

Participants

Ten participants gave their informed consent for the study that was approved by the local ethics committee. All participants (five women, five men) were native Finnish speakers, right-handed (assessed by the Edinburgh handedness

inventory, Oldfield, 1971), with normal (or corrected-to-normal) vision, and had no neurological disorders or diagnosed learning disabilities. Their mean age was 24 years ($SD = 4$ years), and they had either upper secondary or university level education.

The Miniature Language Anigram and the Training Period

Anigram contains 20 nouns (all animal names) and 10 verbs (transitive verbs depicting easily visualized actions), and a limited set of grammatical rules that define a nonadjacent dependency between the agent and patient in a simple active sentence. More specifically, we employed object-marking rules where the sentence object carries one of three suffixes (-s, -r, or -k) determined by the grammatical gender of the sentence subject. Nouns ending with -a/-y are arbitrarily listed as “feminine,” nouns ending with -u/-i “masculine,” and nouns ending with -e/-o are “neutral.” For example, the sentence “the bear hits the cow” is translated “dosuda benosa tunukes” (literally “bear hit cow”) where the feminine subject (ending with -a) determines the object marker -s. This type of object marking was used to avoid transfer effects from the mother tongue, as this specific rule does not occur in Finnish and is also absent in most Indo-European languages (e.g., Swedish, English, German, or French), which are common second languages taught in the Finnish primary school; Finnish children learn at least two secondary languages at school. However, given the presence of object marking and gender as such in many of the languages the participants were familiar with, we cannot completely rule out that some form of transfer may have occurred. The word order was always subject–verb–object, and this was explicitly told to the participants. The images used in the training and in the MEG experiment always depicted the agent on the left and the patient on the right. Each rule appeared with equal probability (33%) both during training and the MEG task.


The miniature language was taught in four daily training sessions (days 1–4), each lasting approximately 1 hr (Figure 1). The experimenter gave explicit information about the syntactical rules at the beginning of each training session. Following this introduction, each rule was exemplified by 30 practice images and the corresponding sentences. Thereafter, the same practice trials, altogether 90, were repeated again but in a random order. Both the example and practice items were black-and-white cartoons of two animal characters that appeared on the screen together with the Anigram sentence. The experimenter first verbally described the pictured event with a Finnish sentence and, subsequently, read out loud the corresponding Anigram sentence. The participant then repeated the Anigram sentence.

After each practice session, learning was tested by naming pairs of animal characters (from pictures without action, i.e., no sentence context), naming verbs (from pictures with only one animal in action), and producing


Training

i. Rule teaching

Explicit rule teaching
Object is marked with -s if subject ends with -a



30 examples




Rules

	subject final	object marker	example
1	-y -a	-s	fituny kimujy garomis the monkey kisses the dog
2	-i -u	-r	garomi lupora madipur the dog points at the lion
3	-e -o	-k	tunuke reduly nadagak the cow pulls the fox

ii. Practice of all rules in random order

90 examples



Teacher says:
"betuga nosepy dosudas"

Student repeats:
"betuga nosepy dosudas"

iii. Evaluation of training

The participant independently produces the learned material

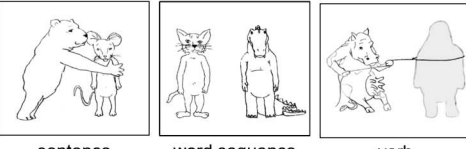


Figure 1. Training period. During each of the four training days, the participant and the instructor completed all three aspects of training. (i) Each of the three rules were explicitly taught to the participant and illustrated with 30 examples. The order in which the rules were presented varied per training day and across participants. (ii) The 90 Anigram sentences encountered during the training session were then presented in a random order. The teacher read out the correct sentence (shown under the image), and the participant repeated it at his or her own pace. (iii) Each training session ended with an evaluation of what had been learned. The participant produced independently full sentences, word sequences (names of two animals), and single verbs, as presented by previously unencountered images.

entire sentences (from pictures of the same type as the practice pictures). The posttraining evaluation was done on a new set of 40 images that were not encountered in the training. On day 4, we additionally tested the participants with 20 completely novel cartoon sentences that had not been used either in the training or in the previous evaluation sessions. Six months after the original experiment, the participants were called back for a behavioral assessment of long-term learning. Participants first attempted to name all nouns and verbs in Anigram from corresponding pictures. They were then asked to produce the correct sentence for a set of 40 completely novel cartoons not used in the previous MEG or

training sessions. If participants were unable to recall the individual depicted words in the last test phase, these were provided in their nominative form to evaluate if a participant nevertheless was able to produce a correct morphosyntactic inflection.

The Finnish language, the participants' mother tongue, was used as the reference condition in the brain imaging part. In Finnish, syntactic processing loads heavily on the morphosyntactic level, whereas word order is of lesser importance. To make the Anigram comparable to Finnish, the syntactic manipulation in this artificial language is similarly contingent on morphosyntactic markers, although the actual morphosyntactic form is notably different. In Finnish total and partial objects (cf. "I drank the milk" vs. "I drank (some) milk") are differentiated through case marking. The verb phrases employed in this study called for partial objects (with the partitive case ending as the object marker: -a, -ä, -ta, or -tä). The most common word order in Finnish is subject-verb-object, which the participants were instructed to use to describe the images.

MEG and Structural MR Measurements

On the day immediately following the end of the training (Day 5), the MEG task (see the description below) was performed both in the native tongue (Finnish) and in the newly learned (Anigram) language, with 240 previously unencountered stimuli. The order of the languages was counterbalanced between the participants in two consecutive MEG measurements on the same day, but so that only one language was used within one measurement session.

During the MEG recording, line drawings of two animal characters were presented on a gray background (Figure 2). The images spanned a visual angle of 4°. There were two main experimental conditions (Sentence, Word sequence) with 120 trials each, presented in a random order. Both conditions consisted of two phases (Preparation phase, Cloze task). The trial began with the presentation of a still image (Preparation phase) with two animals either engaged in an action (Sentence) or passively standing next to each other (Word sequence). The participants were instructed to think about either the depicted sentence (e.g., "the bear begs the mouse") or the word pair (e.g., "the dog, the mouse"), respectively, depending on what was depicted in the image. After 1.5 sec, the task changed into the Cloze task where written words corresponding to the image were shown at the center of the image, one at a time, every 1.5 sec—for the Sentence condition, first the subject, then the verb; for the Word sequence condition, first the name of the animal on the left, then a string of x's to ensure visual stimulation comparable to the sentence condition. However, in place of the final word (sentence object or name of the animal on the right side of the image), a string of question marks appeared, prompting the participant to silently recollect the relevant word. After 1.5 sec, the image disappeared and a single question mark prompted

the participants to name aloud the final word. Before the next trial, a fixation cross appeared for 1 sec. One trial thus lasted for 7 sec in total. Sentence and Word sequence trials were presented in random order, and randomization was done separately for each participant. The MEG data were collected with a 306-channel Neuromag whole-head scanner with a sampling rate of 600 Hz (Elekta Oy, Helsinki) in the MEG Core at the Aalto University.

Structural MR images were acquired with a Signa VH/i 3.0 T MRI scanner (GE Healthcare, Chalfont St Giles, UK) at the Advanced Magnetic Imaging Centre of the Aalto University, using a standard T1-weighted 3-D SPGR sequence.

MEG Data Analysis

The data were band-pass filtered at 0.03–200 Hz. Each step of the task progression was treated as a separate event, that is, one step consisted of either a new picture or a new word/letter string being overlaid on the picture. Average event-related epochs were calculated from 200 before to 1000 msec after each event onset. The 200-msec interval before each event was used as a baseline, and the signal was low-pass filtered at 40 Hz. Artifacts caused by eye movements and blinks were monitored with electrodes placed vertically and horizontally around the eyes and mouth movements with electrodes placed diagonally around the mouth (rejection criterion for both types of electrodes was 150 μ V). Trials containing these artifacts were not included in the analysis. Each condition contained on average 104 ($SD = 16$) accepted

trials. The few incorrect trials (the participants performed close to the ceiling level) were unlikely to affect the averaged response in any relevant way and were not removed from the averages.

As the sensor-level data represent a complex spatial summation of the underlying neural activity, we proceeded to source-level analysis. Coregistration of the MR images to the MEG data followed the Elekta standard procedure and was achieved by determining the head position with respect to the MEG scanner by feeding current to four coils attached to the skin (two on the forehead and one behind each ear). The coil locations were referenced to three landmarks that could be identified from the structural MRIs (left/right preauricular points and nasion). This was done with the help of a 3-D digitizer outside the MEG scanner. 10–20 additional digitized points along the surface of the head were used to aid the alignment with the structural MRI. After manual identification of the fiducial points on the MRI, Elekta software automatically performed the coordinate transformation between the MRI and MEG source spaces.

Distributed source modeling was performed as Minimum Norm Estimates (MNE; “MNE Suite” software package by M. Hämäläinen, Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA; Hämäläinen & Ilmoniemi, 1994). The current distribution that accounted for the measured data and had the minimum power, overall, was constrained to the cortical surface that had been reconstructed from structural MRI data using the Freesurfer software package (Fischl, Liu, & Dale, 2001; Dale, Fischl, & Sereno, 1999). The potential source locations were limited to a grid with 5-mm spacing, and a loose orientation constraint factor

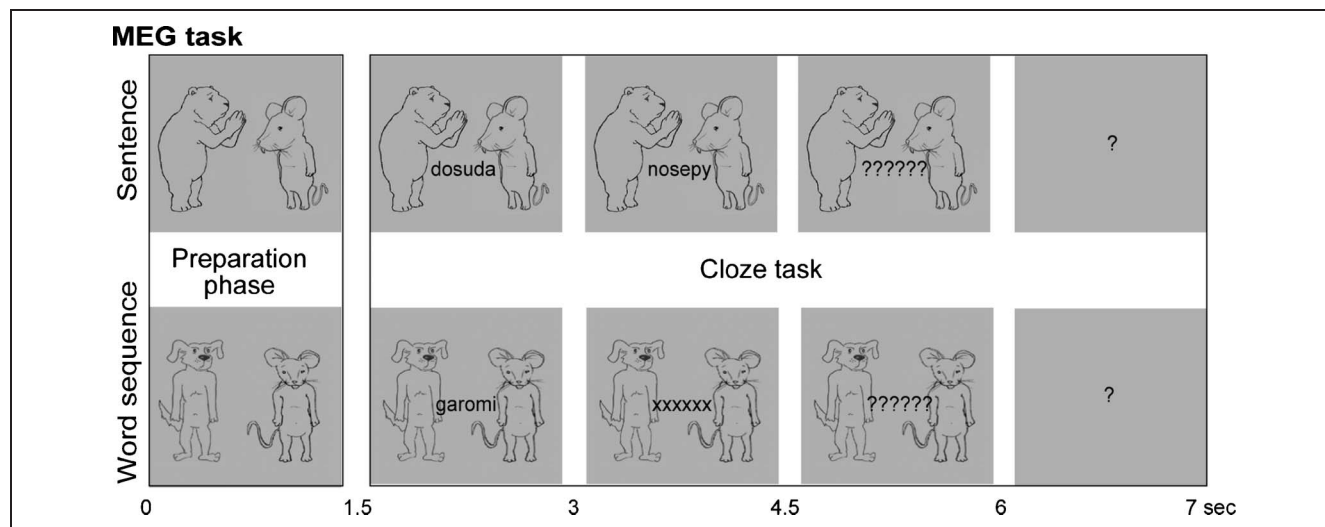
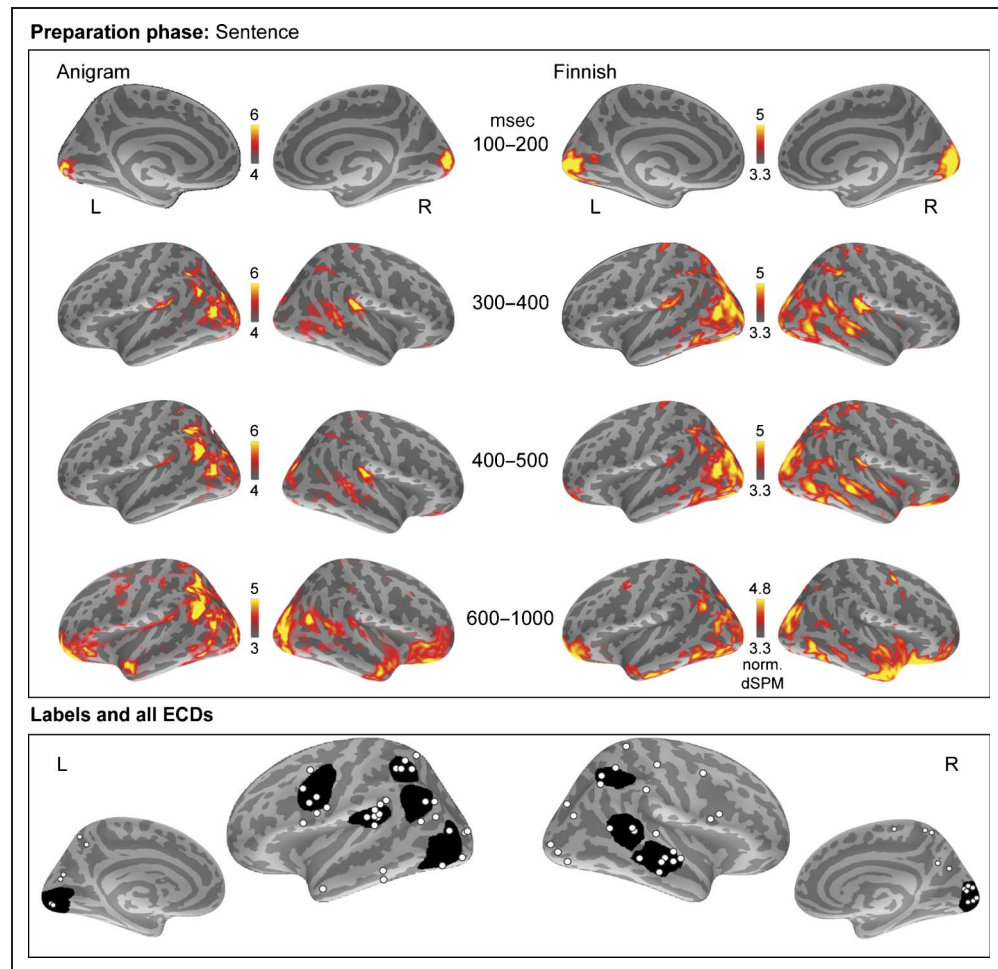


Figure 2. MEG recording. The stimulus images in the MEG task had not been used in the training. In the Preparation phase, the participant was asked to think about the sentence or sequence presented in the image. In the subsequent Cloze task, the corresponding words were superimposed on the image, one by one, and the participant read them silently. The final word was represented by a string of question marks, prompting the participant to name the word silently in its correct form (inflected for sentences, nominative form for word sequences). Overt production of the final word was cued by replacement of the image by a single question mark. The MEG task was performed in both Anigram and the native language (Finnish), with the order of the languages counterbalanced across participants.

Figure 3. Cortical dynamics and ROIs in the Preparation phase. Top: Group-level MNE maps (dSPM normalized within each participant before averaging) integrated within each or the successive time windows for the sentence condition in Anigram (left) and in Finnish (right). Note that the MNEs estimate the loci of activation maxima but do not provide information of the actual shape and extent of the active cortical patch. Bottom: Each white circle represents an ECD of a single participant. Circles overlapping or forming clusters indicate that several participants showed activity in that region. The labels (black) were formed on the basis of the MNE maxima, contingent upon evidence of ECDs (white circles) from at least three participants in the same general area.



of 3.3 was applied to favor currents normal to the cortical surface over transverse ones (Lin, Belliveau, Dale, & Hämmäläinen, 2006). Depth-weighting was used to reduce the bias toward superficial currents inherent to MNE. The forward computation used a single-compartment boundary element model. The results were visualized as noise-normalized MNEs (dynamic SPM, dSPM) that represent the signal-to-noise ratio at each source location as a z score (Dale et al., 2000). The individual dSPM maps were normalized with respect to the maximum level of activity before being morphed onto a standard brain (“fsaverage”) for group-level analysis.

Statistical analyses were performed on the MNE data (absolute activity values, in nAm). The experimental effects on local neural activity were quantified using labels that represented each active region. The labels were identified from the pattern of activity averaged across all the conditions and languages (but separately for the Preparation phase and Cloze task) and covered all separable local field maxima (this approach for determining ROIs has also previously been used by, e.g., Lee, Hämmäläinen, Dyckman, Barton, & Manoach, 2011; Marinkovic et al., 2003). The selection of labels was blind to any differences between the conditions, but we confirmed by visual inspection that

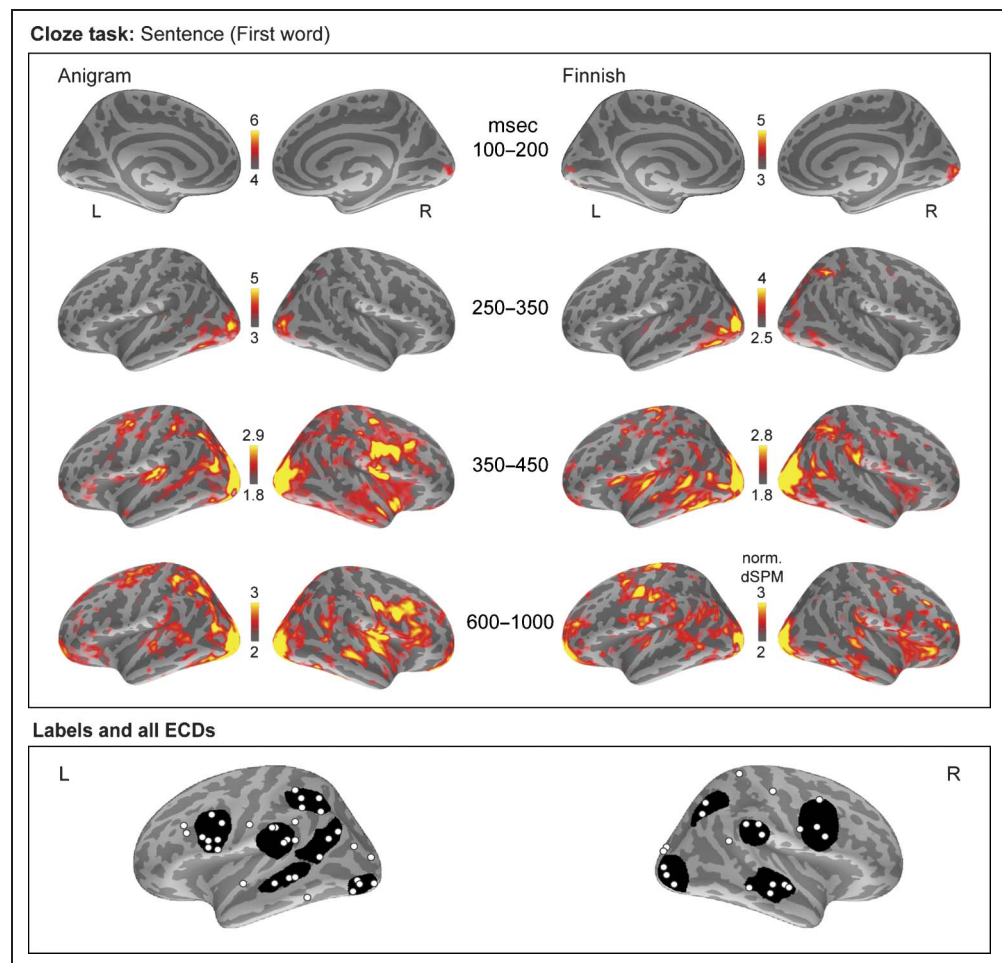
the label selections corresponded well to the activation pattern in each individual condition and that no salient local activation maxima had been overlooked in the label selection (see Figures 3 and 4). We identified 10 labels in the Preparation phase and 11 labels in the Cloze task; these were named according to the approximate anatomical region they coincided with (Preparation phase: left and right medial occipital, left occipito-temporal, left angular, left and right parietal, left posterior peri-sylvian, left frontal, right temporal, and right anterior temporal cortex; Cloze task: left and right occipito-temporal, left posterior temporal, left and right parietal, left and right posterior peri-sylvian, left and right superior temporal, and left and right frontal cortex). It was verified that the labels estimated from the grand average agreed with the pattern observed in each experimental condition and for each language; no additional active areas emerged from this inspection. The shape and location of the labels can be seen in the lower part of Figures 3 and 4; however, because of the field spread in MEG, the exact shape and extent of a label has little influence on the estimated time course of activity. The temporal waveform of activity was extracted from each label by calculating the mean over all vertices included in the label.

As the cortically constrained MNE solution confines all source estimates to the cortical surface, it may map artifact signals to the cortex as well (for example, residues of eye movement or blink artifacts that may remain even after state-of-the-art artifact removal). This potential confound is particularly critical for cortical regions whose activity is inherently hard to reliably detect with MEG, such as the insula and the temporal pole (Hillebrand & Barnes, 2002). We therefore chose to confirm the distributed MNE results by also performing focal source estimation. The dynamics of multiple active cortical areas can be estimated by representing each area as an ECD (Hämäläinen, Hari, Ilmoniemi, Knuutila, & Lounasmaa, 1993). ECD analysis can usually readily pinpoint artifactual sources as they tend to localize outside the brain (e.g., ECD analysis of the residue signals of eye blinks typically localizes the sources next to the eye balls). The ECDs and MNEs represent two extremes of mathematical models that are used to bridge the gap from sensor-level signals to estimates of cortical activity. Both methods yield an estimate of the center of an active area, with the overall appearance of the solution (focal vs. distributed) determined by the model characteristics.

However, neither method provides information about the actual shape or extent of the activated areas.

In the ECD approach, the MEG sensor signals were segregated into separable focal cortical-level spatiotemporal components following well-established analysis procedures (Hari & Salmelin, 2012; Hansen, Kringelbach, & Salmelin, 2010; Salmelin, 2007; Lounasmaa, Hämäläinen, Hari, & Salmelin, 1996). The segregation was performed by means of guided current modeling, where the model parameters of ECD represent the center of an active cortical patch and the mean orientation and strength of electric current within that area (Hämäläinen et al., 1993). The data were scanned visually to find dipolar field patterns, signaling local synchronous neural activation. Each ECD component was determined from a subset of sensors at the point in time when that magnetic field pattern was clearest (Xfit 5.5 software package; Elekta Oy, Helsinki), as described previously by Salmelin (2010). Cortical components explaining at least 85% of the variance of all major deflections of the magnetic field and restricted to the cortical gray matter were included in a multi-ECD model optimized for each participant. The number of ECDs included in these models varied, in the Preparation

Figure 4. Cortical dynamics and ROIs in the Cloze task. Top: Group-level MNE maps (dSPM normalized within each participant before averaging) integrated within each or the successive time windows for the first word in the sentence condition Anigram (left) and Finnish (right). Bottom: Each white circle represents an ECD of a single participant. Circles overlapping or forming clusters indicate that several participants showed activity in that region. The labels (black) were formed on the basis of the MNE maxima, contingent upon evidence of ECDs (white circles) from at least three participants in the same general area.



phase, from 6 to 8 per individual and, in the Cloze task, from 7 to 9. The location and orientation parameters of the ECDs were kept fixed, whereas the strength parameters were allowed to vary to best account for the measured signal from all sensors at each time point. For group-level analysis, the locations of the individual active cortical patches, each represented by their center point, were projected to a common brain space (“fsaverage” of the FreeSurfer package). The MNE and ECD analyses generally displayed good convergence (Figure 3). The MEG analysis was done in accordance with general practice in the field (Gross et al., 2013).

The statistical analyses were done on the MNE result. The activation within each MNE label as a function of time was quantified by the mean amplitudes of sustained responses and peak amplitudes of transient activations. As there is no prior literature on this type of electrophysiological measures for sentence-level speech production, the time windows used in the statistical analysis were selected to incorporate all major deflections observed in the data. In the Preparation phase, the early transient sensory activation in the occipital regions ($n = 2$) was quantified as the peak amplitude and its latency in the time window 0–250 msec. In all other areas ($n = 7$, see Figure 3, bottom), we observed only sustained activity that was quantified by the mean activity at 200–800 msec. In the Cloze task, activity in all areas ($n = 11$) was quantified by the mean amplitude in two windows, at 100–400 msec and 400–1000 msec. We first performed an omnibus across-region repeated-measures ANOVA per time window. This step included three omnibus ANOVAs in the Preparation phase: for both the (i) peak amplitude and (ii) peak latency in the 0–250 msec time window, a region (2) \times language (2) \times context (2) design, and (iii) for the mean amplitude in the 200–800 msec time window, a region (7) \times language (2) \times context (2) design. Analysis of the Cloze task was initiated by omnibus ANOVAs of region (11) \times language (2) \times context (2) \times word presentation order (3), separately for the 100–400 msec and 400–1000 msec time windows. Contingent on discovery of a significant interaction involving region, further analysis was performed per each region. All results were Greenhouse–Geisser corrected when needed. For the sake of brevity, we only report the statistically significant effects. It should also be noted that, as this study utilizes a novel experimental paradigm, the results are exploratory and should be further examined in future work.

RESULTS

Behavioral Results

All participants mastered the miniature language after 4 days of training. They correctly named all nouns (100%, $SD = 0\%$) and virtually all verbs (99.8%, $SD = 0.8\%$) and produced the correct syntactic structure for 98% ($SD =$

2%) of the 40 test sentences. For the 20 test sentences presented only on day 4, the participants scored 96% correct ($SD = 1.5$). During the MEG scan where participants encountered entirely novel stimuli, the rate of correct responses was 95.8% ($SD = 4.2\%$) in Anigram and 99.6% ($SD = 0.4\%$) in Finnish. Although the performance was at ceiling level for both languages, it was nevertheless systematically lower for Anigram (Wilcoxon signed-ranks test; $z = 2.7$, $p < .05$). A behavioral follow-up test at 6 months posttraining revealed that the participants could recall 42% ($SD = 4\%$) of the nouns and 37% ($SD = 2\%$) of the verbs. When testing only for the syntactic structure (three options for case marking), 74% ($SD = 13\%$) of the responses were correct.

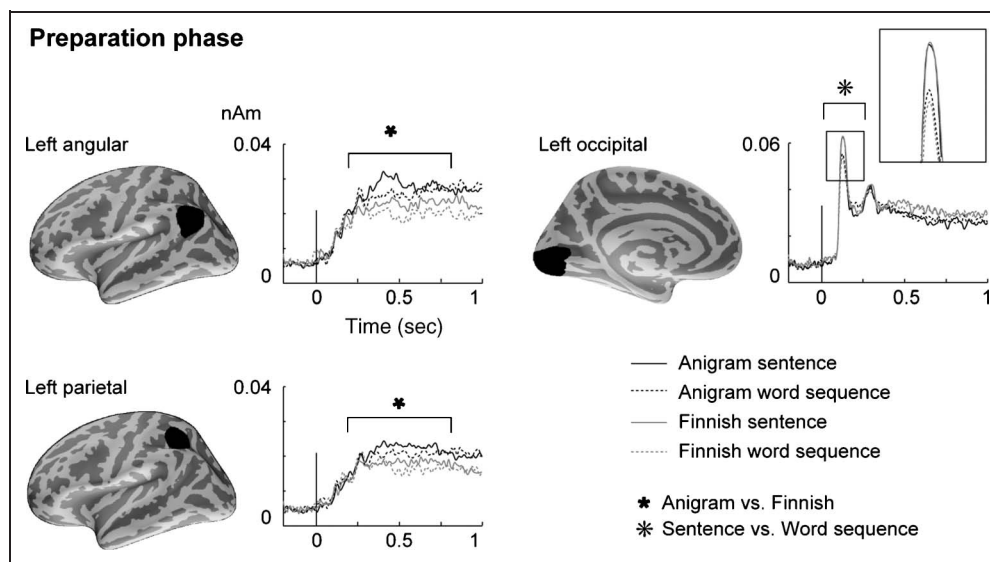
MEG Results

An overview of the sequence of activated areas in the source space (Figures 3 and 4) showed that, as expected, the spatiotemporal configuration of activation maxima was not the same for the two stages of the task (Preparation phase; Cloze task). Accordingly, the quantification using labels and the subsequent statistical testing were done separately for each stage, but based on the average across conditions. The spatial convergence for the distributed MNE and focal ECD source localizations was generally good. As a notable exception, the MNE solution suggested involvement of the anterior–inferior pFC, insula, and temporal pole, whereas the ECD analysis did not show reliable activation in those areas. The lack of convergence between the two complementary source localization methods, low detectability of signals from deep structures such as the insula, and poor sensor coverage over the anterior–inferior part of the brain indicated a low probability of detecting true neural activity from these regions (Hillebrand & Barnes, 2002). Therefore, with an aim to avoid false positives, we opted to not include labels from these regions in the statistical analyses.

Preparation Phase

The only effects of language were observed in the Preparation phase, where participants viewed a picture and prepared for the upcoming sentence or word sequence (Figures 3 and 5). Significant Region \times Language interaction [$F(7, 63) = 2.3$, $p < .05$] in the 200–800 msec time window warranted further analysis within regions, which revealed that activation at in the left angular gyrus and the adjacent parietal cortex was stronger for the newly learned than the native language [left angular: $t(9) = 2.6$, $p < .05$; left parietal: $t(9) = 2.4$, $p < .05$]. At the early time window 0–250 msec, with activity limited to the left and right occipital cortex, the omnibus ANOVA showed a significant Region \times Context interaction [$F(1, 9) = 5.2$, $p < .05$] in the peak amplitude. Within-region pairwise comparison revealed a stronger visual activation (peak

Figure 5. Significant effects in the Preparation phase. Mean time courses of activation in the regions that showed significant effects of Language (Anigram vs. Finnish; bold asterisk) or Context (sentence vs. word sequence; plain asterisk). All statistical analyses were done on the individual nAm values, quantified as the mean amplitude across a time window. The transient occipital activations were additionally quantified by their peak values; in the left medial occipital cortex, a main effect of Context was detected in the peak amplitude.



amplitude) to the images depicting sentences than those related to word sequences in the left occipital cortex [$t(9) = 2.8, p = .02$].

Cloze Task

Significant interactions were observed at 100–400 msec between region, context, and word presentation order [$F(20, 180) = 2.2, p < .01$] as well as between region, context, and language [$F(20, 180) = 1.8, p < .05$], and in the subsequent 400–1000 msec time window between region, language, and word presentation order [$F(20, 180) = 1.8, p < .05$]. Further investigations within each region were thus warranted. The item-by-item progression of the sentence or word sequence decreased the activation in multiple areas in both hemispheres (Figure 6A): at 100–400 msec in the left occipito-temporal [$F(2, 18) = 9.4, p = .01$ (linear trend $p = .01$)], left posterior temporal [$F(2, 18) = 4.2, p = .03$ (quadratic trend $p = .05$)], left posterior peri-sylvian [$F(2, 18) = 5.1, p = .02$ (quadratic trend $p < .01$)], right parietal [$F(2, 18) = 9.9, p = .001$ (linear trend $p = .03$)] and at 400–1000 msec in the left occipito-temporal: $F(2, 18) = 5.1, p = .02$ (linear trend $p = .04$)], left superior temporal [$F(2, 18) = 4.6, p = .02$ (linear trend $p < .01$)] and right occipito-temporal cortex [$F(2, 18) = 12.0, p < .001$ (linear trend $p = .02$)]. In the right superior temporal cortex, the effect was reversed (Figure 6B), that is, activation at 100–400 msec increased with sequence progression, regardless of context or language [$F(2, 18) = 17.0, p < .001$ (linear trend $p < .001$)].

Moreover, the progressive change in activation differed between sentences and word sequences at 100–400 msec in right occipito-temporal cortex [$F(2, 18) = 5.0, p < .001$] and left superior temporal [$F(2, 18) = 6.3, p < .01$]. As illustrated in Figure 6C, in these two areas, the sentence condition demonstrated a salient

quadratic trend, with the weakest response to the second word [right occipito-temporal cortex: $F(1, 9) = 32.8, p < .001$, left superior temporal: $F(1, 9) = 13.2, p < .01$].

No significant modulation was observed for the word sequences, although a decreasing linear trend along the word list approached significance in left superior temporal [$F(1, 9) = 4.7, p = .059$].

DISCUSSION

In this study, we sought to elucidate the cortical mechanisms of foreign language sentence production in adults by comparing a newly learned miniature language (Anigram) with the participants' native tongue (Finnish) and processing of multiword utterances with versus without morphosyntax. The task setup consisted of two phases. First, the participants viewed a cartoon (Preparation phase) and planned for the corresponding output that was either a sentence or a word pair. Second, in the Cloze task, they silently read the initial words/word and then covertly generated the final word either in inflected form (sentence production) or in base form (word pair production). To verify task compliance, the last word was produced overtly after a short delay. The sequential nature of the task made it possible to collect phase-locked signals to the initial picture (Preparation phase) and to each word as well as the covert response (Cloze task). This allowed us to segregate and neurally track different aspects of sentence-level speech production; in natural speech, preparation and production may obviously occur in a more parallel fashion and at a different pace. The task was performed in separate sessions for Anigram and Finnish, but in both cases, the generation of the final word called for retrieval of the corresponding lexical item and the object-marking rule specific to each language.

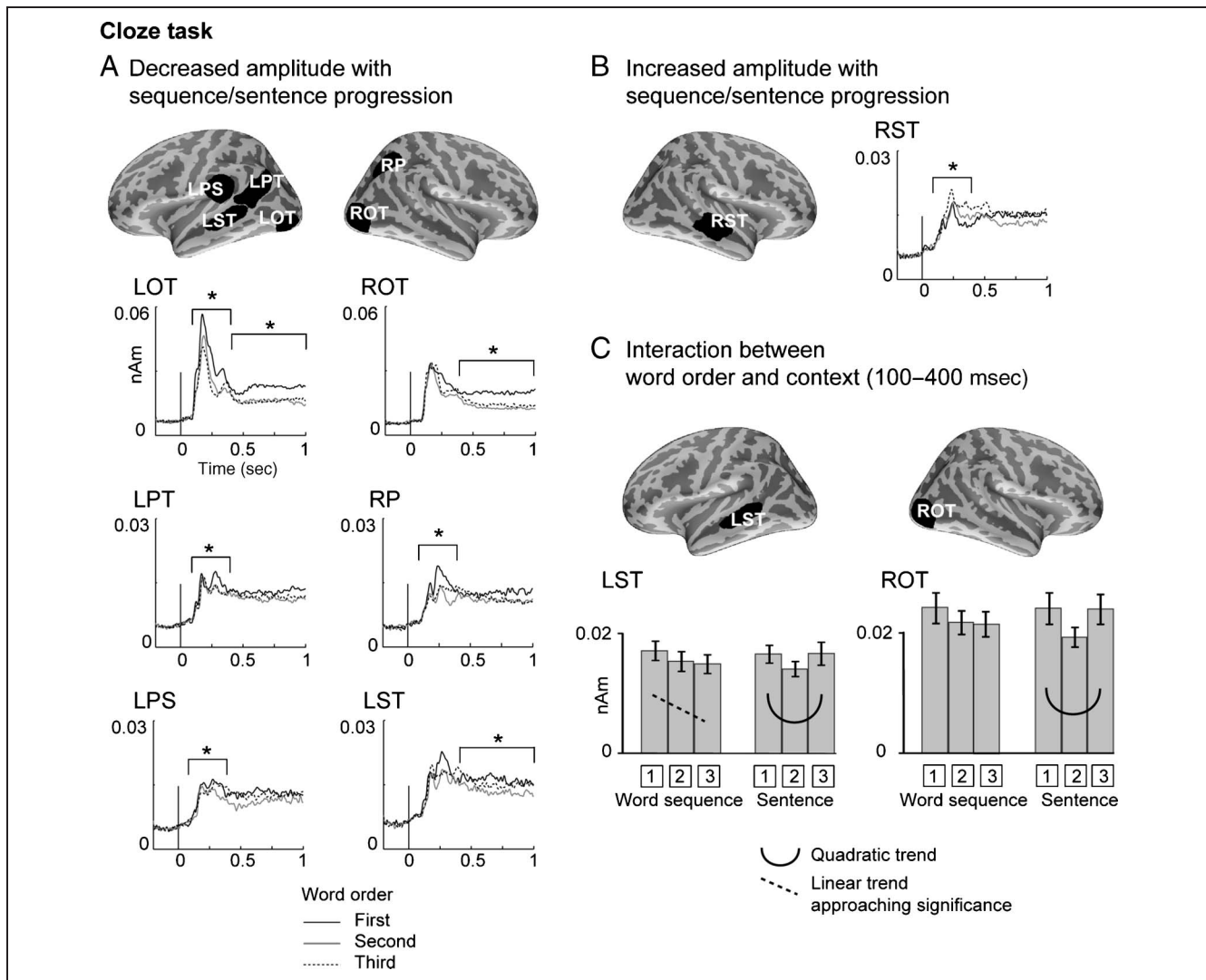


Figure 6. Significant effects in the Cloze task. As there were no significant language effects in this task phase, the results are presented collapsed across languages. (A) Mean time courses of activation in regions that showed a significant main effect of decreasing activation strength with progression along the word sequence or sentence (first, second, third word). (B) Mean time course of activation in the right superior temporal cortex where the amplitude was significantly increased with task progression. (C) Mean activation (\pm SEM) at 100–400 msec in the left superior temporal and right occipito-temporal cortices, the two areas that showed a significant progression-by-context interaction. The amplitudes to the first, second, and third words in each area are plotted separately for word sequences and sentences. Significant linear and quadratic trends are illustrated as lines and curves, respectively (dotted line indicates an effect approaching significance). LOT = left occipito-temporal; LPT = left posterior temporal; LST = left superior temporal; LPS = left posterior peri-sylvian; ROT = right occipito-temporal; RP = right parietal; RST = right superior temporal cortex.

Differences between the Native Language and the Newly Learned Anigram Language

The only difference between the native language and the novel Anigram was observed in the activation strength during the Preparation phase in the left angular gyrus and the adjacent parietal cortex. The Preparation phase assumedly engages thematic role assignment, lexical-semantic access, syntactic sequencing, and STM, as especially the final word is likely to remain active (at least to some degree) until the production prompt. Semantic-conceptual retrieval has been systematically associated with the inferior parietal cortex (Binder, Desai, Graves, & Conant, 2009) as have combinatorial operations

needed in mapping pictorial reference to the semantic contents of a sentence (Bemis & Pylkkanen, 2013; Menenti et al., 2012). The posterior parietal cortex has in turn also been related to working memory functions (Jonides et al., 1998; but see Buchsbaum & D’Esposito, 2008, for a critical perspective). Thus, one possibility is that the elevated sustained amplitudes in the left TPJ (angular gyrus) and parietal cortex for the newly acquired language may be a reflection of more demanding word retrieval and subvocal rehearsal in the newly learned language than in the native tongue. Alternatively, it is not the retrieval per se, but the assignment of each word to the correct thematic role that is the source of this effect. In both cases, a higher challenge in processing

the novel language was supported by the small but consistent differences in behavioral performance between the languages. A third option is that the increased neural effort for Anigram is, in part, because of incidentally activated word representations of either the mother tongue or the other newly learned words. This issue could be evaluated in future speech production studies by utilizing, for example, nontranslatable words or within-language synonyms.

The behavioral results indicated that the participants were in command of the new language after a mere 4 days of training. Nevertheless, they were slightly less proficient in Anigram than in their native language, an observation further strengthened by the fact that the words and grammatical rules were not very well retained at the 6-month follow-up. The average naming performance at the 6-month follow-up was 42% for the nouns in the novel language, comparable to a previous study where French speakers learned a set of 50 English object names and accurately recalled 31% of them 2 months later (Raboyeau et al., 2004). Interestingly, in the study by Hultén et al. (2010), expansion of native vocabulary in a similar fashion resulted in recall levels up to 90% at 2 months postlearning and ~60% at 10 months.

In this study, the influence of language was observed on the neural level as changes of activation strength within the same general cortical areas, in agreement with previous observations (Perani & Abutalebi, 2005). A likely interpretation is that the lower proficiency level in Anigram was associated with increased cognitive effort in task performance, as has been suggested earlier for foreign language processing (Kotz, 2009; Rodríguez-Fornells et al., 2009). The similarity of areas for processing the first (Finnish) and second language (Anigram) might be viewed as an argument against declarative/procedural model (Ullman, 2001). However, a rigorous structural comparison needed for a thorough testing of the declarative/procedural model was not the purpose of this study, nor can MEG spatial estimates of activation patterns (that do not reflect source size or shape, irrespective of the choice of source model) be meaningfully compared by contrasting two conditions. Moreover, recent findings suggest that the degree of native-like processing in the brain is dependent on a number of factors, such as type of exposure and time for consolidation after training (Morgan-Short et al., 2012).

Brain Activations of Multiword Speech Production

Although the learning effects were limited to the Preparation phase, both task stages by themselves also showed effects that are relevant from the perspective of multiword speech production. In general terms, the active areas and their temporal dynamics in the Preparation phase were comparable to those typically observed in picture naming (Vihla et al., 2006; Levelt et al., 1998; Salmelin et al., 1994), whereas the Cloze task part was more reminiscent of that

reported in reading studies (Vartiainen, Aggujaro, et al., 2009; Salmelin, 2007). However, the main stimulus effects in the Preparation phase were found in the angular gyrus, whereas picture naming studies (Vihla et al., 2006; Levelt et al., 1998; Salmelin et al., 1994) tend to report stimulus modulations in somewhat more inferior areas in the posterior superior/middle temporal cortex. This suggests that the Preparation phase of the present task entails more complex or (partly) different type of cognitive processing than basic picture naming, as will be discussed in more detail below.

The left hemisphere regions where activation was modulated as a function of task progression in the Cloze task (but with no concurrent effect of sentence/word pair context) comprised the occipito-temporal, posterior and superior temporal, and posterior peri-sylvian cortex. Their spatiotemporal activation patterns converged with those typical of letter string analysis and lexical access, and their decreasing activity with task progression was in line with that reported for expected sentence endings or semantically related word lists (Vartiainen, Parviainen, & Salmelin, 2009; Helenius, Salmelin, Service, & Connolly, 1998). Similarly, decreasing amplitudes have been reported for confirmed expectation of pseudowords constrained by artificial grammar (Tabullo et al., 2011).

Notably, the modulation of Cloze task progression was reversed in the right temporal cortex where activity increased from the first to the last word, in a linear fashion. Although activation increased similarly for sentences and word pairs, the type of sentences used in this study may have affected the overall processing strategy. In studies of sentence level processing in receptive language, sentences tend to be longer than three words and the content words semantically linked in a fairly realistic manner (e.g., “the man on a vacation lost a bag and a wallet” in Humphries, Binder, Medler, & Liebenthal, 2006). However, the sentences used in production studies are typically much shorter (e.g., Menenti et al., 2012; Indefrey et al., 2001). This is true also for the sentences used here (e.g., “the bear hits the mouse”), which are not only minimal in length but do not have a clear everyday frame of reference. Accordingly, deriving the conceptual meaning of the present animal combinations may be more demanding compared with more conventional sentences with high cloze probabilities. As such, right hemisphere processing may be linked to more demanding thematic role apprehension, a function that has been attributed to the right hemisphere in studies of combinatorial semantics (Graves, Binder, Desai, Conant, & Seidenberg, 2010; Jung-Beeman, 2005) and sentence level production (Menenti et al., 2012; Kircher et al., 2001).

In summary, the present results suggest a functional division of labor between the right and left temporal cortices in multiword speech production. The right temporal cortex displays increasing activity, as the incoming words are linked to the visual scene (word sequences) or the individual items are merged into a meaningful sentence

frame (sentences), whereas the left temporal cortex shows a systematic decrease of activity as the incoming words conform to the internally generated prediction.

Morphosyntactic Processing

Covert production of the final word in the Cloze task facilitated a comparison between morphosyntactically marked object forms and unmarked monomorphemic word forms (and the subsequent overt production ensured task compliance). This comparison revealed a rebound in activation strength for the sentence-final object form in the left superior temporal and right occipito-temporal cortex. Indeed, the left middle superior temporal cortex has been implicated as part of a general sentence production network (Haller et al., 2005). In language comprehension, processing of inflectional morphology has activated very similar areas in the left middle superior temporal cortex (Bölte, Schulz, & Dobel, 2010; Vartiainen, Aggularo, et al., 2009), left posterior middle temporal cortex (Newman, Supalla, Hauser, Newport, & Bavelier, 2010), and right occipito-temporal cortex (Zweig & Pylkkänen, 2009). The present results are also in line with those previous studies and suggest that sustained activity (200–500 msec) in the left superior and middle temporal cortex, typically involved in lexical processing, may also be involved in syntactic processing (Service, Helenius, Maury, & Salmelin, 2007) for both comprehension and production of language.

Comparisons of sentences and sequences in receptive language tasks in West Germanic languages using fMRI have typically reported smaller BOLD activity to word lists compared with sentences in temporal and inferior frontal areas in both hemispheres (Snijders et al., 2009; Humphries et al., 2006). However, as the time resolution of fMRI did not allow for comparison between individual words, those findings are difficult to relate to the present MEG results. The apparent discrepancy with respect to the present results may thus be because of differences in methodologies (time resolution) or, possibly, in tested languages (number of morphosyntactic operations; low in West Germanic languages but high in Finnish).

In this study, the effect for sentence processing emerged in a comparison of two conditions that, as required by the experimental design, were visually different (a cartoon with action vs. two still animals) and demanded retrieval of a different number of words: three in the sentence condition and two in the word pair condition. This naturally warrants caution in the interpretation, but several factors suggest that the effect is not solely related to these stimulus related factors. Although the early visual response of the left occipital cortex was stronger for sentences than for word pair images, this effect was limited to the early visual peak in the Preparation phase and was not present at any later processing stages in the Preparation task. The visual difference between the stimuli was present throughout both task stages, but the differences in the later

Cloze task observed in the right occipito-temporal and left superior temporal cortex were specific only to the morphosyntactic marking (inflectional ending) of the last word. Thus, it seems that the increase of activation to the sentence-final word is not a mere consequence of the image type but suggests that increased processing is needed for application of morphosyntactic object marking by morphosyntactic means.

Methodological Considerations

In this study, both languages activated the left frontal cortex equally for sentences and for word sequences. This finding may seem contradictory to previous fMRI and PET studies on speech production that have linked the left inferior frontal cortex to sentence-level syntactic planning (Golestani et al., 2006; Haller et al., 2005; Indefrey et al., 2001). The discrepancy may be partly a reflection of different experimental designs and baseline conditions across the studies. However, it is also important to note that hemodynamic and electrophysiological measures provide different probes of neural activity. The electrophysiological MEG evoked response is time-locked to the stimulus, whereas the relatively slow hemodynamic fMRI response may, to a greater extent, reflect long-lasting or multiple overlapping cognitive processes. Although MEG and fMRI tend to show involvement of largely comparable brain areas, differences in the relative strength of activation and in the apparent functionality, for example, of the left inferior frontal cortex have been reported (Vartiainen, Liljeström, Koskinen, Renvall, & Salmelin, 2011; Liljeström, Hultén, Parkkonen, & Salmelin, 2009).

As a new experimental paradigm was used in this experiment, possible drawbacks of the design also need to be critically examined. The division into a Preparation phase and a Cloze task may have rendered the task executively more demanding than normal speech. This potential increase in the task demands is, however, the same across all experimental conditions and unlikely to be time-locked to the onset of the stimuli.

Comparing word pairs with sentences presents some additional potential confounds: First, it is possible to perceive the word pair as a sentence, for example, “the mouse stands next to the bear.” However, the participants were not instructed to do this, and furthermore, the Anigram language lacks both the grammar and the vocabulary to formulate such an expression. As the results indicated no differences in word pair processing between Anigram and Finnish, it seems unlikely that the participants perceived the word pairs as sentences. Second, although one can argue that both types of images (or rather, the corresponding names) contain some form of syntactic information, only in the sentence condition was this realized by morphosyntactic composition, which was the topic of our investigation. Third, the presentation of a string of xs instead of a verb makes the word pairs

somewhat artificial, but it also creates a visually comparable yet nonsyntactic condition that was needed as a reference.

Conclusions

The overlap of activation patterns for the novel and the native language indicates that, after only a few days of training, production of the two languages utilizes shared neural resources. However, increased activation of the left parietal cortex and angular gyrus during the initial preparation for a multiword output in a novel language suggests increased cognitive effort as compared with native language processing. Regions associated with morphosyntactic processing in receptive language also seem to be involved when morphosyntactic marking takes place in production. The left and right temporal cortices appear to play different roles in multiword speech production, with the left side involved in prediction of upcoming words and in morphosyntactic marking whereas the right side is engaged in integrating the upcoming words with a particular visual scene or sentence frame.

Acknowledgments

This work was supported by the Academy of Finland (National Centres of Excellence Programme 2006–2011, Neuro2005 Programme, personal grants to R. S. and M. L.), the Sigrid Jusélius Foundation, the Finnish Cultural Foundation, Stiftelsen för Åbo Akademi, and an NOS-HS grant for the Nordic Centre of Excellence in Cognitive Control. The authors would like to express their gratitude to Prof. Elisabet Service for helpful comments on the manuscript and to Prof. Jan-Ola Östman for advice on the linguistic properties on Anigram.

Reprint requests should be sent to Annika Hultén, Brain Research Unit, O.V. Lounasmaa Laboratory, Aalto University, P.O. Box 15100, 00076 Aalto, Finland, or via e-mail: annika.hulten@aalto.fi.

REFERENCES

- Awad, M., Warren, J. E., Scott, S. K., Turkheimer, F. E., & Wise, R. J. (2007). A common system for the comprehension and production of narrative speech. *Journal of Neuroscience*, *27*, 11455–11464.
- Bemis, D. K., & Pylkkanen, L. (2013). Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. *Cerebral Cortex*, *23*, 1859–1873.
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*, 2767–2796.
- Blank, S. C., Scott, S. K., Murphy, K., Warburton, E., & Wise, R. J. (2002). Speech production: Wernicke, Broca and beyond. *Brain*, *125*, 1829–1838.
- Bölte, J., Schulz, C., & Dobel, C. (2010). Processing of existing, synonymous, and anomalous German derived adjectives: An MEG study. *Neuroscience Letters*, *469*, 107–111.
- Brownsett, S. L., & Wise, R. J. (2010). The contribution of the parietal lobes to speaking and writing. *Cerebral Cortex*, *20*, 517–523.
- Buchsbaum, B. R., & D'Esposito, M. (2008). The search for the phonological store: From loop to convolution. *Journal of Cognitive Neuroscience*, *20*, 762–778.
- Clark, E. V., & Hecht, B. F. (1983). Comprehension, production, and language acquisition. *Annual Review of Psychology*, *34*, 325–349.
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage*, *9*, 179–194.
- Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., et al. (2000). Dynamic statistical parametric mapping: Combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron*, *26*, 55–67.
- Davis, M. H., Di Betta, A. M., Macdonald, M. J., & Gaskell, M. G. (2009). Learning and consolidation of novel spoken words. *Journal of Cognitive Neuroscience*, *21*, 803–820.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*, 283–321.
- Dhanjal, N. S., Handunnethi, L., Patel, M. C., & Wise, R. J. (2008). Perceptual systems controlling speech production. *Journal of Neuroscience*, *28*, 9969–9975.
- Dobel, C., Gumnior, H., Bölte, J., & Zwitserlood, P. (2007). Describing scenes hardly seen. *Acta Psychologica*, *125*, 129–143.
- Fischl, B., Liu, A., & Dale, A. M. (2001). Automated manifold surgery: Constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Transactions on Medical Imaging*, *20*, 70–80.
- Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences, U.S.A.*, *99*, 529–534.
- Golestani, N., Alario, F. X., Meriaux, S., Le Bihan, D., Dehaene, S., & Pallier, C. (2006). Syntax production in bilinguals. *Neuropsychologia*, *44*, 1029–1040.
- Goodglass, H., & Kaplan, E. (1983). *The assessment of aphasia and related disorders* (2nd ed.). Philadelphia, PA: Lea & Febiger.
- Graves, W. W., Binder, J. R., Desai, R. H., Conant, L. L., & Seidenberg, M. S. (2010). Neural correlates of implicit and explicit combinatorial semantic processing. *Neuroimage*, *53*, 638–646.
- Grönholm, P., Rinne, J. O., Vorobyev, V., & Laine, M. (2005). Naming of newly learned objects: A PET activation study. *Brain Research, Cognitive Brain Research*, *25*, 359–371.
- Gross, J., Baillet, S., Barnes, G. R., Henson, R. N., Hillebrand, A., Jensen, O., et al. (2013). Good practice for conducting and reporting MEG research. *Neuroimage*, *65*, 349–363.
- Haller, S., Radue, E. W., Erb, M., Grodd, W., & Kircher, T. (2005). Overt sentence production in event-related fMRI. *Neuropsychologia*, *43*, 807–814.
- Hämäläinen, M. S., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography—Theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, *65*, 413–497.
- Hämäläinen, M. S., & Ilmoniemi, R. J. (1994). Interpreting magnetic fields of the brain: Minimum norm estimates. *Medical & Biological Engineering & Computing*, *32*, 35–42.
- Hansen, P. C., Kringelbach, M. L., & Salmelin, R. (2010). *MEG an introduction to methods*. New York: Oxford University Press.

- Hari, R., & Salmelin, R. (2012). Magnetoencephalography: From SQUIDs to neuroscience: Neuroimage 20th Anniversary Special Edition. *Neuroimage*, *61*, 386–396.
- Helenius, P., Salmelin, R., Service, E., & Connolly, J. F. (1998). Distinct time courses of word and context comprehension in the left temporal cortex. *Brain*, *121*, 1133–1142.
- Hillebrand, A., & Barnes, G. R. (2002). A quantitative assessment of the sensitivity of whole-head MEG to activity in the adult human cortex. *Neuroimage*, *16*, 638–650.
- Hultén, A., Laaksonen, H., Vihla, M., Laine, M., & Salmelin, R. (2010). Modulation of brain activity after learning predicts long-term memory for words. *Journal of Neuroscience*, *30*, 15160–15164.
- Hultén, A., Vihla, M., Laine, M., & Salmelin, R. (2009). Accessing newly learned names and meanings in the native language. *Human Brain Mapping*, *30*, 976–989.
- Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Journal of Cognitive Neuroscience*, *18*, 665–679.
- Indefrey, P., Brown, C. M., Hellwig, F., Amunts, K., Herzog, H., Seitz, R. J., et al. (2001). A neural correlate of syntactic encoding during speech production. *Proceedings of the National Academy of Sciences, U.S.A.*, *98*, 5933–5936.
- Jonides, J., Schumacher, E. H., Smith, E. E., Koeppe, R. A., Awh, E., Reuter-Lorenz, P. A., et al. (1998). The role of parietal cortex in verbal working memory. *Journal of Neuroscience*, *18*, 5026–5034.
- Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences*, *9*, 512–518.
- Kim, K. H., Relkin, N. R., Lee, K. M., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature*, *388*, 171–174.
- Kircher, T. T., Brammer, M., Tous Andreu, N., Williams, S. C., & McGuire, P. K. (2001). Engagement of right temporal cortex during processing of linguistic context. *Neuropsychologia*, *39*, 798–809.
- Kircher, T. T., Brammer, M. J., Williams, S. C., & McGuire, P. K. (2000). Lexical retrieval during fluent speech production: An fMRI study. *NeuroReport*, *11*, 4093–4096.
- Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain and Language*, *109*, 68–74.
- Lee, A. K., Hämäläinen, M. S., Dyckman, K. A., Barton, J. J., & Manoach, D. S. (2011). Saccadic preparation in the frontal eye field is modulated by distinct trial history effects as revealed by magnetoencephalography. *Cerebral Cortex*, *21*, 245–253.
- Levelt, W. J. M., Praamstra, P., Meyer, A. S., Helenius, P., & Salmelin, R. (1998). An MEG study of picture naming. *Journal of Cognitive Neuroscience*, *10*, 553–567.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–38; discussion 38–75.
- Liljeström, M., Hultén, A., Parkkonen, L., & Salmelin, R. (2009). Comparing MEG and fMRI views to naming actions and objects. *Human Brain Mapping*, *30*, 1845–1856.
- Lin, F. H., Belliveau, J. W., Dale, A. M., & Hämäläinen, M. S. (2006). Distributed current estimates using cortical orientation constraints. *Human Brain Mapping*, *27*, 1–13.
- Lounasmaa, O. V., Hämäläinen, M., Hari, R., & Salmelin, R. (1996). Information processing in the human brain: Magnetoencephalographic approach. *Proceedings of the National Academy of Sciences, U.S.A.*, *93*, 8809–8815.
- Marinkovic, K., Dhond, R. P., Dale, A. M., Glessner, M., Carr, V., & Halgren, E. (2003). Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron*, *38*, 487–497.
- Menenti, L., Gierhan, S. M., Segaert, K., & Hagoort, P. (2011). Shared language: Overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. *Psychological Science*, *22*, 1173–1182.
- Menenti, L., Petersson, K. M., & Hagoort, P. (2012). From reference to sense: How the brain encodes meaning for speaking. *Frontiers in Psychology*, *2*, 384.
- Morgan-Short, K., Finger, I., Grey, S., & Ullman, M. T. (2012). Second language processing shows increased native-like neural responses after months of no exposure. *PLoS One*, *7*, e32974.
- Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2011). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience*, *24*, 933–947.
- Newman, A. J., Supalla, T., Hauser, P., Newport, E. L., & Bavelier, D. (2010). Dissociating neural subsystems for grammar by contrasting word order and inflection. *Proceedings of the National Academy of Sciences, U.S.A.*, *107*, 7539–7544.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.
- Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, *15*, 202–206.
- Price, C. J. (2010). The anatomy of language: A review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences*, *1191*, 62–88.
- Raboyeau, G., Marie, N., Balduyck, S., Gros, H., Demonet, J.-F., & Cardebat, D. (2004). Lexical learning of the English language: A PET study in healthy French subjects. *Neuroimage*, *22*, 1808–1818.
- Rodríguez-Fornells, A., Cunillera, T., Mestres-Missé, A., & de Diego-Balaguer, R. (2009). Neurophysiological mechanisms involved in language learning in adults. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *364*, 3711–3735.
- Salmelin, R. (2007). Clinical neurophysiology of language: The MEG approach. *Clinical Neurophysiology*, *118*, 237–254.
- Salmelin, R. (2010). Multi-dipole modeling in MEG. In P. C. Hansen, M. L. Kringelbach, & R. Salmelin (Eds.), *MEG an introduction to methods* (pp. 124–155). New York: Oxford University Press.
- Salmelin, R., Hari, R., Lounasmaa, O. V., & Sams, M. (1994). Dynamics of brain activation during picture naming. *Nature*, *368*, 463–465.
- Service, E., Helenius, P., Maury, S., & Salmelin, R. (2007). Localization of syntactic and semantic brain responses using magnetoencephalography. *Journal of Cognitive Neuroscience*, *19*, 1193–1205.
- Snijders, T. M., Vosse, T., Kempen, G., Van Berkum, J. J., Petersson, K. M., & Hagoort, P. (2009). Retrieval and unification of syntactic structure in sentence comprehension: An fMRI study using word-category ambiguity. *Cerebral Cortex*, *19*, 1493–1503.
- Tabullo, A., Sevilla, Y., Pasqualetti, G., Vernis, S., Segura, E., Yorio, A., et al. (2011). Expectancy modulates a late positive ERP in an artificial grammar task. *Brain Research*, *1373*, 131–143.
- Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism-Language and Cognition*, *4*, 105–122.

- Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92*, 231–270.
- Vartiainen, J., Aggajaro, S., Lehtonen, M., Hultén, A., Laine, M., & Salmelin, R. (2009). Neural dynamics of reading morphologically complex words. *Neuroimage, 47*, 2064–2072.
- Vartiainen, J., Liljeström, M., Koskinen, M., Renvall, H., & Salmelin, R. (2011). fMRI BOLD signal and MEG evoked responses yield different neural functionality in reading. *Journal of Neuroscience, 31*, 1048–1058.
- Vartiainen, J., Parviainen, T., & Salmelin, R. (2009). Spatiotemporal convergence of semantic processing in reading and speech perception. *Journal of Neuroscience, 29*, 9271–9280.
- Vihla, M., Laine, M., & Salmelin, R. (2006). Cortical dynamics of visual/semantic vs. phonological analysis in picture confrontation. *Neuroimage, 33*, 732–738.
- Zweig, E., & Pykkänen, L. (2009). A visual M170 effect of morphological complexity. *Language and Cognitive Processes, 24*, 412–439.