Brain & Language 122 (2012) 71–80

Contents lists available at SciVerse ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

The neuronal infrastructure of speaking

Laura Menenti [a,b,*], Katrien Segaert [a,c], Peter Hagoort [a,c]

[a] Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands
[b] Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, United Kingdom
[c] Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Accepted 23 April 2012
Available online 19 June 2012

Keywords:
Language production
fMRI
fMRI adaptation
Repetition suppression
Semantics
Syntax
Lexical processing

ABSTRACT

Models of speaking distinguish producing meaning, words and syntax as three different linguistic components of speaking. Nevertheless, little is known about the brain's integrated neuronal infrastructure for speech production. We investigated semantic, lexical and syntactic aspects of speaking using fMRI. In a picture description task, we manipulated repetition of sentence meaning, words, and syntax separately. By investigating brain areas showing response adaptation to repetition of each of these sentence properties, we disentangle the neuronal infrastructure for these processes. We demonstrate that semantic, lexical and syntactic processes are carried out in partly overlapping and partly distinct brain networks and show that the classic left-hemispheric dominance for language is present for syntax but not semantics.
© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Twas brillig and the slithy toves
Did gyre and gimble in the wabe;
All mimsy were the borogoves
And the mome raths outgrabe.(Carroll, 1871)

What are toves? No clue. Or, well…? They are things that can be plural, be 'slithy', and can 'gyre' and 'gimble'. We are apparently able to process grammar in the absence of meaningful words. Indeed, models of speech production (Garrett, 1975; Levelt, 1989; Levelt, Roelofs, & Meyer, 1999; Vigliocco & Hartsuiker, 2002) agree that what we say (semantics) and how we say it (syntax) are different aspects of the speaking process. In this study, we investigate whether the neuronal infrastructure underlying speaking also shows this distinction.

Speaking is the conversion of an intention to communicate a message into a linearized string of speech sounds. An essential step in this process is the retrieval of the relevant concepts and the specification of semantic structure. One key aspect of semantics is *thematic role structure*. It refers to the relation between the different concepts and events, or 'who does what to whom'. Schematically one can state this as a predicate with arguments: BITE(-DOG, MAN) means there is a BITE event, a DOG is the agent of this event (the one who bites), while MAN is the patient (the one who is bitten).

The thematic role structure BITE(DOG, MAN) can be expressed in a variety of ways, depending on the choice of a syntactic structure: "The man was bitten by a dog", "The dog bit a man", " It was a dog that bit the man", or even "Did the dog bite a man?". Of course, though these sentences share the same thematic role structure, they do have subtle differences in meaning. In this paper we will focus on thematic role structure as a key aspect of semantic structure. Words play a central role in connecting semantic and syntactic structure (Hagoort, 2005; Levelt, 1989; Vosse & Kempen, 2000). The mental lexicon, our memory for language, contains information on semantic, syntactic and phonological properties of words. When we prepare an utterance, the relevant concepts (TO BITE, DOG, MAN) are retrieved from memory. The syntactic properties (e.g., word class, grammatical gender) of the associated words are in turn also activated. For example, the concept BITE belongs to a lexical item 'to bite' that has the syntactic property of being a verb. This verb takes a subject and an object, which defines the specifics of the sentence structure that the verb can enter into. For example, the sentence "The dog bites a man to the woman" is ungrammatical because the verb 'bites' cannot take an indirect object. Through unification of these syntactic constraints of activated lexical items a syntactic structure for the sentence is generated.

* Corresponding author. Address: Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands.
E-mail address: laura.menenti@mpi.nl (L. Menenti).

0093-934X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bandl.2012.04.012

Though cognitive models specify the components of speech production in much detail, very little is known about the neuronal infrastructure for the linguistic processes involved in producing multi-word utterances. So far cognitive neuroimaging research on language has focused on comprehension (Gernsbacher & Kashak, 2003), and the few studies on language production are mostly either on single word production (Alario, Chainay, Lehericy, & Cohen, 2006; Karbe, Herholz, Weber-Luxenburger, Ghaemi, & Heiss, 1998; Kircher, Brammer, Tous Andreu, Williams, & McGuire, 2001; Tremblay & Gracco, 2006; Tremblay & Gracco, 2010; Tremblay & Small, 2011b; Wise et al., 2001; Zheng, Munhall, & Johnsrude, 2010; for a review, see Indefrey & Levelt, 2004), or on covert production (den Ouden, Hoogduin, Stowe, & Bastiaanse, 2008). The neuroimaging studies that have investigated overt sentence-level production either treat sentence production as a unitary process (Awad, Warren, Scott, Turkheimer, & Wise, 2007; Blank, Scott, Murphy, Warburton, & Wise, 2002; Brownsett & Wise, 2010; Foki, Gartus, Gesissler, & Beisteiner, 2008; Kemeny, Ye, Birn, & Braun, 2005; Kircher, Brammer, Williams, & McGuire, 2000; Stephens, Silbert, & Hasson, 2010), or isolate only one component of speech production (Haller, Radue, Erb, Grodd, & Kircher, 2005; Indefrey et al., 2001; Kircher, Oh, Brammer, & McGuire, 2005; Tremblay & Small, 2011a). It is, therefore, unknown to what degree the different cognitive stages in speech production also recruit different neuronal networks.

In a recent functional Magnetic Resonance Imaging (fMRI) study, we have disentangled semantic, lexical, and syntactic processes during sentence production and comprehension by using a fMRI adaptation paradigm (Menenti, Gierhan, Segaert, & Hagoort, 2011). fMRI adaptation is a phenomenon in which the Blood Oxygen Level Dependent (BOLD) response in neuronal populations sensitive to a stimulus is affected after repetition of that stimulus (Grill-Spector, Henson, & Martin, 2006; Krekelberg, Boynton, & Wezel, 2006). fMRI adaptation can also be used to identify areas sensitive to particular stimulus attributes, by manipulating repetition of different attributes independently. Importantly, in such a paradigm multiple simultaneously occurring stimulus properties can be spatially segregated.

In Menenti et al. (2011) we applied this logic to investigate the global overlap between speaking and listening in three important components of language: semantic, lexical and syntactic processing. By comparing fMRI adaptation effects for semantic, lexical, and syntactic repetition in speaking and listening, we found that for all three components, the neuronal infrastructure was largely shared between speaking and listening. Bilateral posterior middle temporal gyri were involved in sentence-level semantic processing. Left posterior and anterior middle temporal gyrus, and left inferior and middle frontal gyrus, and the homologous areas on the right, were involved in lexical processing. Left posterior middle temporal gyrus and left inferior frontal gyrus were involved in syntactic processing.

In that study, however, we left unanswered the question of exactly how all the areas we found involved conspire in producing, or understanding, an utterance. For comprehension, this issue has been extensively addressed in the literature (Bookheimer, 2002; Friederici, Ruschemeyer, Hahne, & Fiebach, 2003; Hagoort, 2005; Martin, 2003). For production, however, our study provided the first useful set of data to look at the interplay of these components in more detail.

In this paper, we investigate the neuronal infrastructure for different steps in speech production. We look at overlap and segregation of semantic, lexical and syntactic processing in the brain, to investigate how the whole process of speaking, from thought to spoken words, may be performed through the cooperation of a network of brain areas.

In our study, 20 Dutch participants described pictures by producing short sentences, while lying in the MRI-scanner. Syntactic, semantic or lexical aspects of the spoken sentences could be either repeated or novel between two subsequent sentences (Fig. 1). Syntactic repetition is manipulated independently of both lexical and semantic repetition, enabling us to find areas uniquely sensitive to syntax. Lexical and sentence-level semantic repetition cannot be manipulated orthogonally to each other, as it is hard to communicate the same message without, at least partly, using the same words. Our design, however, allows us to manipulate repetition of word meaning and thematic role structure respectively, while keeping the other constant. This allows us to distinguish areas sensitive to word meaning from areas involved in sentence-level semantic processing. To counter MR-artefacts due to speaking, we used an fMRI-sequence that increases sensitivity and reduces motion artefacts (Buur, Poser, & Norris, 2009; Poser, Versluis, Hoogduin, & Norris, 2006).

2. Methods

For experimental methods, also see Menenti et al. (2011).

2.1. Subjects

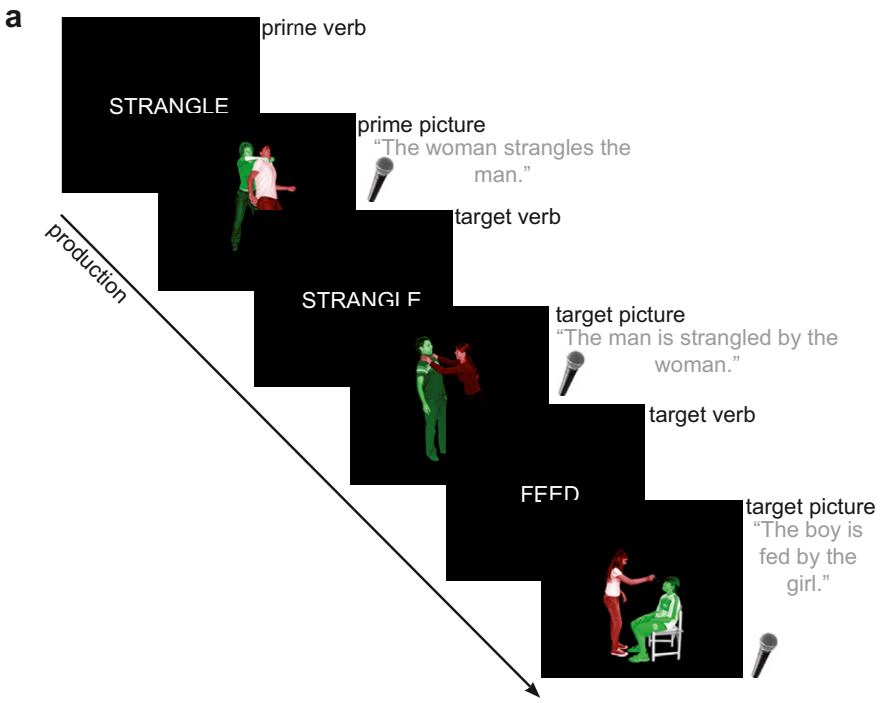
Twenty-four (seven male) healthy right-handed (as assessed through an adapted version of the Edinburgh Handedness Inventory (Oldfield, 1971)) Dutch native speakers with normal or corrected to normal vision (mean age 22 years, range 20–28) participated in the experiment. Four subjects were excluded from analysis (one male) due to technical problems with the data.

2.2. Stimuli

Our target stimuli were photographs that depicted 36 transitive events such as *kiss*, *help*, *strangle* with the agent (doer) and patient (undergoer) of this action. These pictures could be described with transitive sentences, which were spoken by the participants. The pictures were displayed with one actor colored in red and the other in green, to cue the participants in our *stoplight paradigm* (see below). Besides the target stimuli, we also had filler pictures that depicted either intransitive or locative events.

2.3. Design

The design is illustrated in Fig. 1. There were three factors (syntax, semantics, words), with two levels (repeated/novel) each. The syntactic structure of subsequent sentences could either be the same (e.g., active–active) or different (e.g., active–passive). Separate repetition of meaning (semantic repetition) and individual words (lexical repetition) allowed us to distinguish areas that are sensitive to sentence meaning, and those that are sensitive to repetition of words but not to repetition of sentence meaning. We could not do so by using orthogonal factors, since it is not possible to repeat sentence meaning without, at least partly, repeating words. This effectively led to a $2 \times 2 \times 2$ design (semantics repeated/novel, words repeated/novel, syntax repeated/novel) with two empty cells (semantics repeated – words novel, for both novel and repeated syntax). We analyzed the effect of semantic repetition only in those conditions where the lexical content was repeated (thereby keeping the factor words constant), and the effect of word repetition only for those conditions where semantic structure was novel (therefore, keeping the factor meaning constant). Syntactic repetition was orthogonal to these factors.


a
prirne verb
STRANGLE
prime picture
"The woman strangles the man."
target verb
STRANGLE
target picture
"The man is strangled by the woman."
target verb
FEED
target picture
"The boy is fed by the girl."
production

b
prime trial
sentence (spoken or heard):
The woman strangles the man.
syntax: active
words: {man, woman, strangle}
semantics: STRANGLE(WOMAN,MAN)
target trials
novel syntax
repeated syntax
novel words
The boy is fed by the girl.
passive
{boy,girl,feed}
FEED(GIRL,BOY)
The girl feeds the boy.
active
{boy,girl,feed}
FEED(GIRL,BOY)
repeated words
novel semantics
The woman is strangled by the man.
passive
{man,woman,strangle}
STRANGLE (MAN,WOMAN)
The man strangles the woman.
active
{man,woman,strangle}
STRANGLE (MAN,WOMAN)
repeated semantics
The man is strangled by the woman.
passive
{man, woman, strangle}
STRANGLE (WOMAN,MAN)
The woman strangles the man.
active
{man, woman, strangle}
STRANGLE (WOMAN,MAN)

Fig. 1. Design and stimuli. (A) Procedure. Subjects produced descriptions of scenes, which consisted of short sentences containing action (e.g., strangle), agent (woman) and patient (man). We guided their production through the *stoplight paradigm*. Subjects were instructed to use the verb that preceded the picture and to start their description with the green character. In effect, this forced them to produce either an active ('The woman strangles the man.') or a passive sentence ('The man is strangled by the woman.') Each target served as prime for the subsequent target. (B) Design. In the figure, the crucial relationship is between the prime picture (top) and the target. Six prime-target combinations were used. Semantic, lexical, and syntactic content could be repeated (e.g., for syntax: active–active or passive–passive) or novel (active–passive or passive–active) between two subsequent sentences. Syntactic repetition could be orthogonally manipulated with respect to semantic and lexical repetition. Semantic and lexical repetition were each compared between conditions where the other factor was constant. For example, *The woman is strangled by the man* and *The man is strangled by the woman* both contain the same words as the prime *The woman strangles the man*. For the effect of semantic repetition, we tested for differences between only these two conditions (STRANGLE(MAN, WOMAN)–STRANGLE(MAN, WOMAN) vs. STRANGLE(MAN, WOMAN)–STRANGLE(WOMAN, MAN)), leaving the condition with novel words (STRANGLE(MAN, WOMAN)–FEED(BOY, GIRL)) out of the comparison. The same logic applies to the effect of lexical repetition, where we left the condition with repeated thematic role structure (STRANGLE(MAN, WOMAN)–STRANGLE(MAN, WOMAN)) out of the comparison.

2.4. Task

Each target picture was preceded by its corresponding verb. Participants were instructed to describe the picture with a short sentence, using the presented verb. Each sentence had to start with the green person or object. In this sentence they had to mention the green person or object before the red person or object. We call this the *stoplight paradigm* (see Fig. 1A (Menenti et al., 2011)).

2.5. Procedure

The outline of the experiment was, therefore, as follows: Subjects produced a sentence for a list of pictures, consisting of alternating mini-blocks of fillers and targets. The list started with 3–6 filler trials. These were followed by 3–6 transitive trials. The first of the sentences was considered a prime, followed by 2–5 target sentences, each also serving as primes for the next target sentence. Each picture was always preceded by its corresponding verb. Between a prime sentence and its target sentence, the semantic, lexical or syntactic content could be repeated or could be novel, leading to six different prime-target combinations (see Fig. 1B). Repetition of semantic content consisted of two sentences with the same thematic role structure (e.g., STRANGLE(MAN, WOMAN)). Lexical repetition consisted of two sentences containing the same words (e.g. {man, woman, strangle}). If the syntactic structure was repeated, then the two sentences had the same syntactic structure (e.g., passive). After the 3–6 transitive pictures, another 3–6 filler pictures followed, and this continued for a total of 64 blocks, split over two runs. Each trial consisted of the following events: first, the verb appeared for 500 ms. After an ISI of 500–1500 ms the picture appeared for 2000 ms. After another ISI of 4000–7000 ms the next trial started. The experiment consisted of two runs of 47 min. Between the two runs the participants underwent an anatomical T1 scan and had a short break outside the MRI-scanner. Participants completed a short practice block in the scanner before the actual experiment started.

2.6. Data acquisition and analysis

2.6.1. Behavioral data

The reaction times for correct responses to the target items were analysed in a repeated measures ANOVA using SPSS, as were the error data. To separate participants' speech from the scanner sound and extract reaction times, the speech recordings were bandpass filtered with a frequency band of 250–4000 Hz and smoothed with a width half the sampling rate. Speech onsets were determined and linked to the stimulus presentation times to extract the reaction times. We discarded from the RT-data all trials with incorrect responses, all trials with an RT below 300 ms or above 3000 ms and all trials more than 3SD's from each individual subject's mean. These trials were analyzed for the error data analysis. From the fMRI data we discarded all incorrect responses.

2.6.2. fMRI data

Data acquisition took place in a 3-T Siemens Magnetom Tim-Trio magnetic resonance imaging scanner using a 12-channel surface coil. To acquire functional data, we used parallel-acquired inhomogeneity-desensitized fMRI (Poser et al., 2006). In this multiecho-planar imaging sequence, images are acquired at multiple echo times following a single excitation. Accelerated parallel imaging reduces motion and susceptibility artifacts and thus is a good method to acquire data when participants are speaking in the scanner.

For the functional scans, the TR was 2398 ms and each volume consisted of 31 slices of 3 mm thickness with slice-gap of 17% of the slice thickness. The voxel size was $3.5 \times 3.5 \times 3$ mm and the field of view was 224 mm. Functional scans were acquired at multiple TEs following a single excitation (TE1 at 9.4 ms, TE2 at 21.2 ms, TE3 at 33 ms, TE4 at 45 ms, and TE5 at 56 ms with echo spacing of 0.5 ms) so that there was a broadened $T2^*$ coverage. The number of slices did not allow for complete full brain coverage in most participants. The experimenter always made sure that the frontal and temporal lobes, where activations of interest were expected, were always included. This meant that in many participants data from the top of the head was not acquired. A whole-brain high resolution structural T1-weigthed MPRAGE sequence was performed to characterize participants' anatomy (TR = 2300 ms, TE = 3.03 ms, 192 slices with voxel size of 1 mm^3, FOV = 256), accelerated with GRAPPA parallel imaging (Griswold et al., 2002).

The fMRI data were preprocessed and analyzed with SPM5 (Friston et al., 1995). The first five images were discarded to allow for T1 equilibration. Then the five echoes of the remaining images were realigned to correct for motion artefacts (estimation of the realignment parameters is done for one echo and then copied to the other echoes). The five echoes were combined into one image with a method designed to filter task-correlated motion out of the signal (Buur et al., 2009). First, echo two to five (i.e. TE2, TE3, TE4 and TE5) were combined using a weighting vector dependent on the measured differential contrast to noise ratio. The time course of an image acquired at a very short echo time (i.e. TE1) was used as a voxelwise regressor in a linear regression for the combined image of TE2, TE3, TE4 and TE5. Weighting of echoes was calculated based on 25 volumes acquired before the actual experiment started. The resulting images were coregistered to the participants' anatomical scan, normalized to MNI space and spatially smoothed using a 3D isotropic Gaussian smoothing kernel (FWHM = 8 mm).

Our $2 \times 2 \times 2$ design with the factors semantic, lexical and syntactic repetition resulted in six conditions, as two cells were empty (see Section 2.3). These were crossed with two additional factors: whether the target structure was an active or a passive and whether the picture depicted adults or children. This led to 24 main regressors (six conditions $\times$ two syntactic structures $\times$ two

sets of actors). Separate regressors were included in the first level model for the verbs, fillers, prime items and incorrect responses. Note that incorrect responses were put in a separate regressor and therefore excluded from analysis. All events were modeled as stick functions with a duration of 0 ms. These regressors were all convolved with the canonical hemodynamic response function provided by SPM5. Finally, the realignment parameters were added as regressors of no interest. For the second level model we used the beta-images of the 24 main regressors. The second level contrasts were computed pooling over syntactic structure and actors, as these factors do not pertain to the questions at hand in the current manuscript. The cluster size was used as the test statistic and only clusters significant at $P < .05$ corrected for multiple non-independent comparisons are reported. Local maxima are also reported for all clusters with their respective family-wise error (FWE) corrected *P*-values. Anatomical labels are derived from Brodmann's atlas and the Automatic Anatomical Labeling Atlas (Tzourio-Mazoyer et al., 2002), both as included in the software package MRIcron. These atlases associate every anatomical area to a set of MNI coordinates, and hence produce a deterministic approximation of anatomical locations.

We computed two types of additional analyses on the data reported in Menenti et al. (2011). To find areas sensitive to different components of speaking, and where we can therefore postulate the relevant processes to interface, we used conjunction analyses. To find areas which are selectively sensitive to only one component, and which may therefore be involved in one specific processing step, we performed exclusive masking. We did this by computing exclusive masks at $p < .05$ uncorrected. Note that due to the very nature of the type of statistical framework we employ, we cannot prove that an effect does *not* exist. However, if an effect does not survive tresholding at $p < .05$ uncorrected voxelwise, it may be said to be very weak and unreliable at the very least.

3. Results

3.1. Behavioral results

The behavioral data are reported in Fig. 2. Repeated measures of the error data revealed only an effect of lexical repetition ($F(1,19) = 7.56$, $p < .013$). Repeated measures GLM's on the reaction time data revealed additive, non-interacting effects of semantic ($F = 4.6$, $p < .045$), lexical ($F = 57.582$, $p < .000$) and syntactic repetition ($F = 5.277$, $p < .033$). For both error data and reaction time data there were no interactions between syntax and semantics, and between syntax and words. The interaction between words and semantics cannot be tested, as each of the factors is tested at only one level of the other (e.g., the difference between semantically novel and semantically repeated sentences is tested only for those conditions where words are repeated).

3.2. fMRI results

The main effects of the three factors are reported in Fig. 3 and Table 1, the specific analyses of overlap and segregation between semantics, words and syntax are reported in Fig. 4 and Table 2. As the latter are a more specific and hence informative version of the main effects, we focus on those in what follows.

Specifying relational semantic (e.g., thematic role) structure, lexical retrieval, and syntactic encoding are three distinct operations in speaking. It is, therefore, to be expected that at least some areas in the brain are sensitive to only one of these components of sentence production. We first tested each of these contributions separately, while excluding the other effects through exclusive masking. The only area sensitive to the repetition of thematic role structure without showing either a syntactic or lexical effect is the right middle temporal gyrus (MTG; BA 37 and 21). Repetition of words (with all their associated properties) elicits suppression in left anterior MTG (BA 21 and 22), and in a medial cluster in the left precuneus (BA 29 and 30). Repetition of syntactic structure leads to suppression in the left inferior frontal gyrus (LIFG; BA 44/45/47), left precentral gyrus (BA 44/9/6), and the left inferior parietal lobule (IPL; BA 40/19/7).

Since the lexicon plays a crucial role as interface between syntax and semantics, we expected to find areas that are sensitive to all three of the core properties. A cluster extending from left MTG to left ITG (BA 37) shows suppression to syntactic, semantic and lexical repetition.

Syntactic encoding is achieved through the unification of lexical-syntactic information into syntactic structures. This should be reflected in areas sensitive to both lexical and syntactic repetition while not showing a semantic effect. A cluster extending from the left MTG to ITG (BA 37/21/20), and the left precentral gyrus (BA 6) showed this pattern. There were no areas sensitive to both lexical and semantic but not syntactic repetition.

Finally, besides the lexicon, higher-level biases (e.g., the tendency to express a transitive event as an active sentence) could interface between semantic and syntactic processes. Potential areas involved in these high-level matching procedures should be sensitive to syntactic and semantic, but not lexical repetition. There is one such area in the left superior parietal cortex (superior parietal lobule, precuneus; BA 7).

4. Discussion

This study provides the first integrated detailed picture of the neuronal infrastructure underlying linguistic processes in speaking.

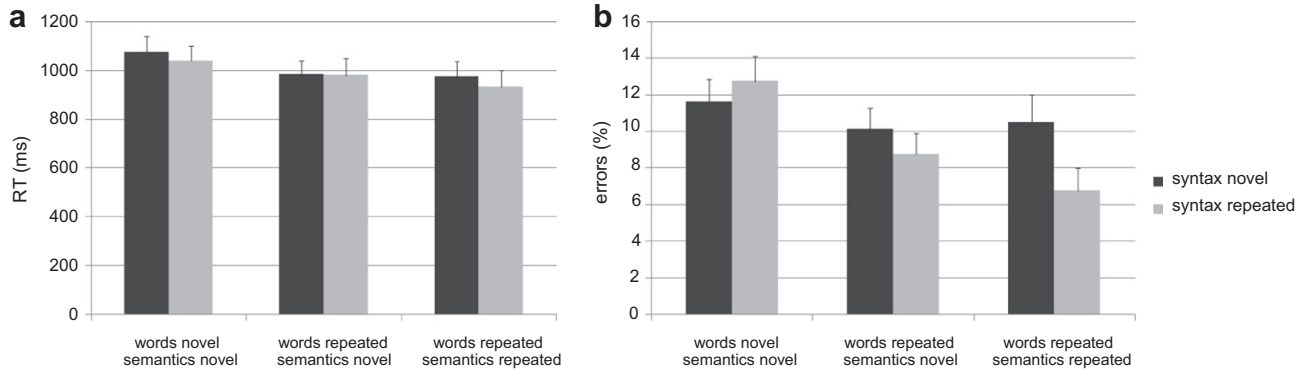


Fig. 2. Behavioral data. (A) Reaction times. Repeated measures GLM's revealed additive, non-interacting effects of semantic, lexical and syntactic repetition. (B) Error data. Repeated measures GLM's revealed only an effect of lexical repetition.

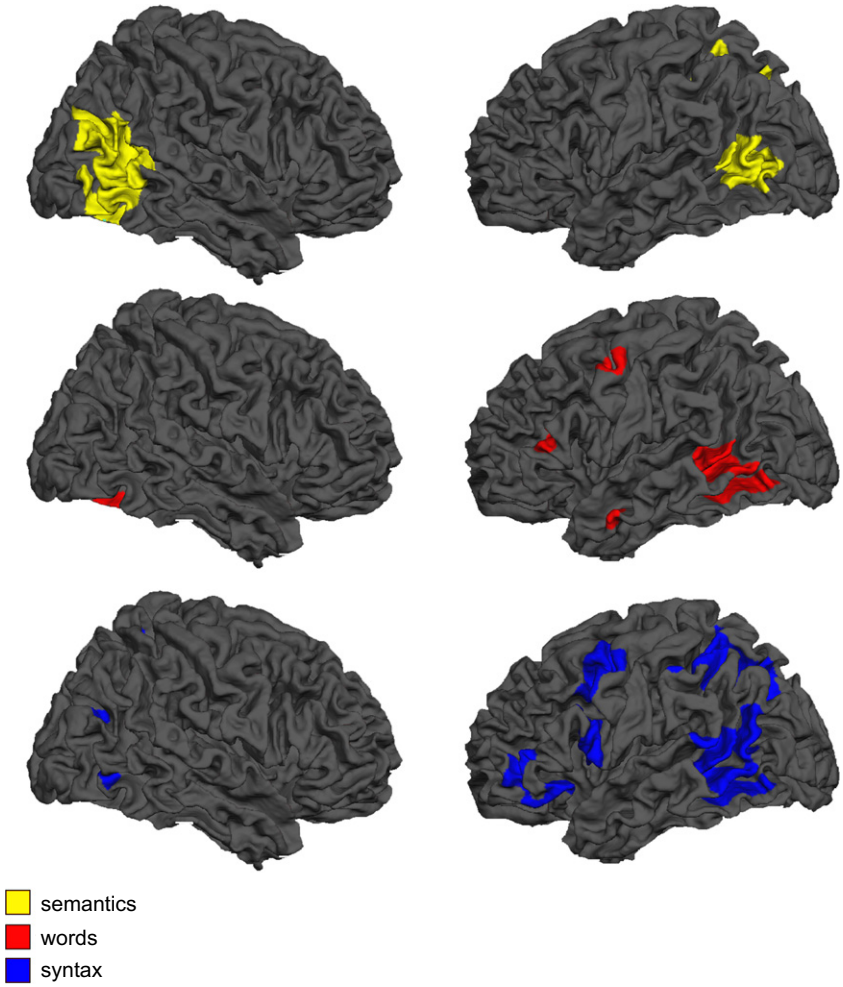


Fig. 3. Main effects of semantic, lexical and syntactic repetition. Shown are all areas that show an effect of either semantic, lexical or syntactic repetition, but with no restriction on sensitivity to the other effects. The effects displayed in Fig. 4 are a subset of these. All effects $p < .05$ corrected at cluster level (threshold $p < .0001$ uncorrected voxel-wise). See also Table 1.

Our data disclose the following organizational principles: (i) Core operations in speaking, such as thematic role specification, lexical retrieval, and syntactic encoding, are in part subserved by non-overlapping brain areas. (ii) These unique areas are embedded in a network of additional areas that provide the necessary neuronal infrastructure for interfacing between the semantic, lexical, and syntactic operations during speaking. (iii) There is a clear gradient in the asymmetry between the two hemispheres, with the process of syntactic encoding lateralized to the left perisylvian cortex, while the specification of the relational content of an utterance relies on contributions of the right hemisphere. Moreover, different parts of the left temporal lobe play a crucial role in the storage of word-related information. Note that all these areas are most likely not language-specific, but in the context of speaking they play their own, relatively unique, role.

Our results provide a clear picture of the brain areas involved in speaking. If we progress through the linguistic steps in producing a sentence, the speaker's first step is to generate a semantic (thematic role) structure that can be syntactically encoded such as BITE(DOG, MAN). The input to this process is the non-linguistic mental representation the speaker wants to talk about. We postulate the right posterior MTG (sensitive to repetition of sentence-level semantic structure but not of individual concepts) to be relevant for providing this input to the generation of thematic role structure. The bilateral MTG's, sensitive to repetition of thematic role structure but also to some (weak) extent to lexical semantics, are presumably involved in generating such a thematic role structure.

Once the speaker's intention has led to the generation of a semantic structure, this structure contains the semantic properties of the words to be inserted. Retrieval of the relevant lemmas further triggers their corresponding syntactic properties, which are then encoded into a syntactic structure for the sentence. The lexicon is the crucial interface between semantic and syntactic processes and the left posterior MTG is the prime candidate area contributing to this interfacing in the lexicon, as it shows sensitivity to lexical but also semantic and syntactic processes. This is in accordance with models postulating a crucial role of posterior temporal areas in storage of word knowledge (Hagoort, 2005).

The anterior MTG, however, is sensitive to repetition of words but not to any sentence-level properties. In the present study repetition of words conflates the repetition of many different properties (semantic, syntactic, and phonological) of these words. Our results, therefore, cannot tell us whether the lexical repetition effect is specific to one of these properties, or common to all of them. In a meta-analysis on word production, the anterior MTG was found to be commonly activated during picture and word naming and therefore postulated to be involved in the selection of lexical concepts (Indefrey & Levelt, 2004), an idea that our results corroborate. This same area is also affected in semantic dementia (Wilson et al., 2009), a progressive neurodegenerative disorder mainly

Table 1
Main effects of semantic, lexical and syntactic repetition. Listed are the MNI-coordinates for the first three local maxima for each significant cluster in the relevant comparisons ($p < .05$ corrected cluster-level, threshold $p < .0001$ uncorrected voxel-wise). Anatomical labels are derived from the Automatic Anatomical Labeling atlas (Tzourio-Mazoyer et al., 2002) and from Brodmann's atlas. Cluster-level statistics are listed for each cluster, voxel-level statistics also for local maxima.

Effect	Cluster	BA	Anatomical label	Global and local maxima			Cluster-level		Voxel-level	
				x	*y*	*z*	*K*	*p*	*T*	*p (FWE)*
Semantics	1	37	L middle temporal gyrus	*−48*	*−66*	*6*	843	0.000	5.89	0.000
		37	n/a	−38	−58	14			4.29	0.205
	2	37	R middle temporal gyrus	56	−66	0	1457	0.000	5.58	0.001
		39	R middle temporal gyrus	*44*	*−64*	*20*			5.52	0.001
		37	R middle temporal gyrus	52	−56	6			4.80	0.031
	3	5	R precuneus	6	−56	54	303	0.000	4.50	0.098
		7	L precuneus	*−8*	*−62*	*52*			4.31	0.195
		7	L superior parietal lobule	−20	−68	50			4.17	0.299
Words	1	30	L precuneus	−4	−50	16	2118	0.000	7.51	0.000
		29	R precuneus	*8*	*−50*	*14*			5.75	0.000
		n/a	Cerebellum	*−6*	*−44*	*0*			5.48	0.001
	2	20	L inferior temporal gyrus	−46	−50	−14	1341	0.000	7.09	0.000
		37	L fusiform gyrus	−38	−38	−20			5.40	0.002
		37	n/a	*−46*	*−52*	*2*			5.33	0.003
	3	37	R fusiform gyrus	36	−56	−20	148	0.003	5.67	0.001
	4	10	L superior frontal gyrus, medial	−6	58	30	74	0.027	5.01	0.013
	5	6	L precentral gyrus	−50	0	48	79	0.02	4.53	0.09
	6	11	L superior frontal gyrus, orbital	−2	54	−8	122	0.01	4.49	0.104
		11	L rectus	−2	42	−18			3.86	0.635
Syntax	1	6	L precentral gyrus	*−42*	*2*	*48*	951	0.000	6.07	0.000
		6	L precentral gyrus	−38	−2	56			5.52	0.001
		44	L precentral gyrus	−50	8	34			4.97	0.015
	2	37	L inferior occipital gyrus	*−46*	*−60*	*−12*	1175	0.000	5.85	0.000
		37	L inferior temporal gyrus	−42	−46	−12			5.40	0.002
		21	L middle temporal gyrus	−58	−54	8			5.10	0.009
	3	40	L inferior parietal gyrus	*−36*	*−52*	*54*	1892	0.000	5.37	0.003
		40	L inferior parietal gyrus	−42	−44	50			5.33	0.003
		7	L superior parietal lobule	−24	−72	42			5.25	0.004
	4	45	L inferior frontal gyrus p triangularis	*−44*	*42*	*0*	93	0.014	4.26	0.228
		45	L inferior frontal gyrus p triangularis	−42	44	10			4.05	0.419

Table 2
Results. Listed are the MNI-coordinates for the first three local maxima for each significant cluster in the relevant comparisons ($p < .05$ corrected cluster-level, threshold $p < .001$ uncorrected voxel-wise). Anatomical labels are derived from the Automatic Anatomical Labeling atlas(Tzourio-Mazoyer et al., 2002) and from Brodmann's atlas, both in MRIcron. Cluster-level statistics are listed for each cluster, voxel-level statistics also for local maxima.

Effect	Cluster	BA	Anatomical label	Global and local maxima			Cluster-level		Voxel-level	
				x	*y*	*z*	*K*	*p*	*T*	*p (FWE)*
Only semantics	1	37	R middle temporal gyrus	*42*	*−62*	*16*	367	0.003	5.11	0.008
		21	R middle temporal gyrus	52	−46	4			4.64	0.047
		39	R middle occipital gyrus	56	−66	26			4.45	0.098
Only words	1	29	L precuneus	*−8*	*−50*	*14*	1428	0.000	6.36	0.000
		30	L precuneus	2	−48	16			6.36	0.000
		27	L lingual gyrus	−6	−44	4			5.31	0.003
	2	22	L middle temporal gyrus	*−52*	*−12*	*−8*	244	0.003	4.36	0.163
		21	L middle temporal gyrus	−54	−2	−22			4.30	0.197
		21	L middle temporal gyrus	−58	−4	−14			3.77	0.745
Only syntax	1	40	L inferior parietal lobule	*−40*	*−48*	*50*	732	0.000	5.13	0.007
		19	L middle occipital gyrus	−30	−66	36			5.02	0.012
		7	L inferior parietal lobule	−34	−58	48			4.63	0.060
	2	44	L precentral gyrus	*−50*	*8*	*34*	647	0.000	4.97	0.015
		9	L precentral gyrus	−42	8	48			4.87	0.024
		6	L precentral gyrus	−36	0	58			4.82	0.029
	3	45	L inferior frontal gyrus, pars triangularis	*−44*	*42*	*0*	361	0.004	4.26	0.228
		45	L inferior frontal gyrus, pars triangularis	−42	44	10			4.05	0.419
		47	L inferior frontal gyrus, pars orbitalis	−42	38	−18			3.78	0.729
Syntax + words, no semantics	1	37	L inferior temporal gyrus	*−44*	*−44*	*−10*	257	0.017	4.59	0.072
		21	L middle temporal gyrus	−56	−46	4			4.27	0.010
		20	L inferior temporal gyrus	−46	−46	−18			3.94	0.018
	2	6	L precentral gyrus	*−50*	*0*	*48*	192	0.051	4.53	0.006
Syntax + semantics, no words	1	7	L superior parietal lobule	*−20*	*−68*	*50*	707	0.000	4.17	0.299
		7	L superior parietal lobule	−32	−52	58			4.13	0.338
		7	L precuneus	−10	−64	56			3.83	0.675
Syntax, words and semantics	1	37	L middle temporal gyrus	*−56*	*−56*	*4*	604	0.000	4.32	0.188
		37	L middle temporal gyrus	−56	−66	−2			3.99	0.486
		37	L inferior occipital gyrus	−46	−58	−12			3.89	0.564

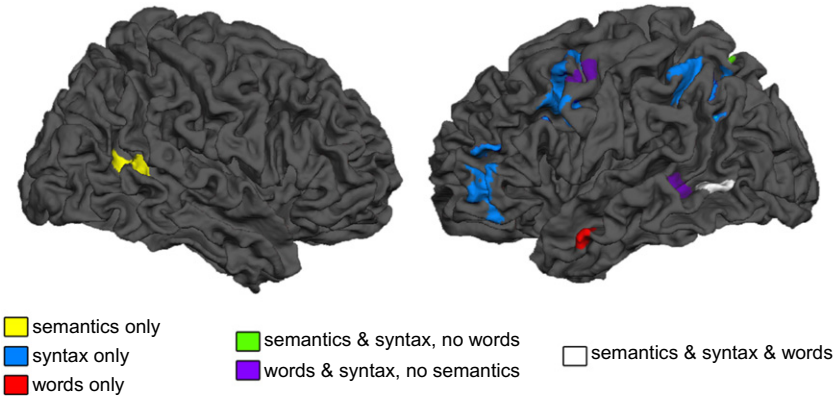


Fig. 4. Results. Areas sensitive to only one factor (semantics, words or syntax), two of these (semantics and syntax, or words and syntax), and the one area sensitive to all three factors. All effects $p < .05$ corrected at cluster level (threshold $p < .001$ uncorrected voxel-wise). For results excluding one or two factors, the effects are masked exclusively at $p < .05$ uncorrected. See also Table 2.

characterized by loss of the ability to remember meanings of words. It has, therefore, been proposed as an amodal hub in semantic memory for concepts (Patterson, Nestor, & Rogers, 2007).

Processing of meaning at the word- and sentence level clearly produces different suppression patterns: sentence-level semantics is processed by the bilateral MTG's, but only the left MTG is sensitive to lexical repetition. Our results also clarify the role of right hemisphere areas in semantic processing in healthy subjects: the right MTG focuses on relational meaning beyond the single concept level. This result is consistent with the idea that the right hemisphere is relatively more involved in computing (non-linguistic) situation models that, in speaking, provide the input to specifying the propositional content of an utterance (Bookheimer, 2002; Long, Baynes, & Prat, 2005; Mitchell & Crow, 2005). In a follow-up study we have disentangled linguistic and non-linguistic semantic repetition in a similar paradigm, and shown that the rMTG sensitivity to semantic repetition is non-linguistic in nature (Menenti, Petersson, & Hagoort, 2012).

After conceptual encoding, the speaker formulates a syntactic structure. These two processes are necessarily intertwined: though different syntactic structures can express the same thematic role structure, they do carry subtle differences in meaning – which is why the speaker might consider different syntactic structures for one and the same thematic role structure in the first place. The message BITE(MAN, DOG) can be expressed as *The dog was bitten by a man* or *The man bit a dog*. We normally take the first sentence to be about the dog, and to focus on that it was a man (and not a snake) that bit the dog, while the second sentence is about the man and tells us he bit a dog (and not a sandwich).

Irrespective of the chosen syntactic structure, syntactic encoding is thought to occur through the unification of syntactic frames (Hagoort, 2005; Levelt, 1989; Vosse & Kempen, 2000). The LIFG is involved in generating a syntactic structure: left BA 44 has previously been found involved in syntactic encoding in production in healthy subjects (Haller et al., 2005; Indefrey et al., 2001), along with lower BA 6. These regions have also been found to show syntactic priming in comprehension (Weber & Indefrey, 2009) and to be sensitive to syntactic complexity (Meltzer, McArdle, Schafer, & Braun, 2010).

The precentral gyrus showed suppression for syntactic repetition and lexical repetition combined. In a study investigating comprehension of syntactic structures of varying syntactic complexity in semantically reversible and irreversible sentences, this region was found to be sensitive to syntactic complexity only for reversible sentences (Meltzer et al., 2010). The authors interpreted this as additional cognitive resources strategically recruited to compute syntactic structures under high cognitive demands. As we assume, both lexical frames and syntactic unification processes are involved in syntactic encoding, repeating both from one trial to the next presumably reduces cognitive demands.

There were also areas showing only sensitivity to syntactic repetition. These non-lexical areas suggest the additional involvement of more general procedures in syntactic encoding.

Besides the lexicon, more general interfacing mechanisms between semantics and syntax, such as the tendency to make the agent (DOG) the subject of the sentence (producing "*The dog bites the man*" rather than "*The man is bitten by the dog*"), may exist. The superior parietal lobule (sensitive to syntax and semantics but not to words) might play a role in this type of non-lexical matching processes.

5. Conclusions

In sum, in an fMRI experiment that imaged the speaking brain we investigated three core aspects of speaking. We have shown that speaking relies on both areas with more specialized and with relatively global sensitivities to semantic, syntactic and lexical information. In addition, the language dominance of the left hemisphere shows up most clearly for syntax, but appears weaker for relational semantics.

Acknowledgments

This research was supported by the NWO Spinoza Prize awarded to P.H. We are very grateful to Erno Hermans, Peter Koopmans, Pieter Buur and Lennart Verhagen for invaluable help in setting up and analyzing the study. We thank the many friends, colleagues and family members who posed for our stimuli.

References

Alario, F. X., Chainay, H., Lehericy, S. p., & Cohen, L. (2006). The role of the supplementary motor area (SMA) in word production. *Brain Research, 1076*(1), 129–143. http://dx.doi.org/10.1016/j.brainres.2005.11.104.

Awad, M., Warren, J. E., Scott, S. K., Turkheimer, F. E., & Wise, R. J. S. (2007). A common system for the comprehension and production of narrative speech. *Journal of Neuroscience, 27*(43), 11455–11464.

Blank, S. C., Scott, S. K., Murphy, K., Warburton, E., & Wise, R. J. S. (2002). Speech production: Wernicke, broca and beyond. *Brain, 125*(8), 1829–1838. http://dx.doi.org/10.1093/brain/awf191.

Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual review of neuroscience, 25*, 151–188.

Brownsett, S. L. E., & Wise, R. J. S. (2010). The contribution of the parietal lobes to speaking and writing. *Cerebral Cortex, 20*(3), 517–523. http://dx.doi.org/10.1093/cercor/bhp120.

Buur, P. F., Poser, B. A., & Norris, D. G. (2009). A dual echo approach to removing motion artefacts in fMRI time series. *Magnetic Resonance in Medicine, 22*(5), 551–560.

Carroll, L. (1871). Through the looking-glass, and what Alice found there.

den Ouden, D.-B., Hoogduin, H., Stowe, L. A., & Bastiaanse, R. (2008). Neural correlates of Dutch Verb Second in speech production. *Brain and Language, 104*(2), 122–131.

Foki, T., Gartus, A., Gesissler, A., & Beisteiner, R. (2008). Probing overtly spoken language at sentential level—A comprehensive high-field BOLD–fMRI protocol reflecting everyday language demands. *NeuroImage, 39*, 1613–1624.

Friederici, A. D., Ruschemeyer, S.-A., Hahne, A., & Fiebach, C. J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: Localizing syntactic and semantic processes. *Cerebral Cortex, 13*(2), 170–177.

Friston, K. J., Holmes, A., Worsley, K., Poline, J.-B., Frith, C., & Frackowiak, R. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping, 2*, 189–210.

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 133–177). New York: Academic Press.

Gernsbacher, M. A., & Kashak, M. P. (2003). Neuroimaging studies of language production and comprehension. *Annual Review of Psychology, 54*, 91–114.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences, 10*(1), 14–23.

Griswold, M. A., Jakob, P. M., Heidemann, R. M., Nittka, M., Jellus, V., Wang, J., et al. (2002). Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magnetic Resonance in Medicine, 47*(6), 1202–1210.

Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences, 9*(9), 416–423.

Haller, S., Radue, E. W., Erb, M., Grodd, W., & Kircher, T. (2005). Overt sentence production in event-related fMRI. *Neuropsychologia, 43*(5), 807–814. http://dx.doi.org/10.1016/j.neuropsychologia.2004.09.007.

Indefrey, P., Brown, C. M., Hellwig, F., Amunts, K., Herzog, H., Seltz, R. J., et al. (2001). A neural correlate of syntactic encoding during speech production. *PNAS, 98*(10), 5933–5936.

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*, 101–144.

Karbe, H., Herholz, K., Weber-Luxenburger, G., Ghaemi, M., & Heiss, W.-D. (1998). Cerebral networks and functional brain asymmetry: Evidence from regional metabolic changes during word repetition. *Brain and Language, 63*(1), 108–121. http://dx.doi.org/10.1006/brln.1997.1937.

Kemeny, S., Ye, F. Q., Birn, R., & Braun, A. R. (2005). Comparison of continuous overt speech fMRI using BOLD and Arterial Spin Labeling. *Human Brain Mapping, 24*, 173–183.

Kircher, T. T. J., Brammer, M., Tous Andreu, N., Williams, S. C. R., & McGuire, P. K. (2001). Engagement of right temporal cortex during processing of linguistic context. *Neuropsychologia, 39*(8), 798–809. http://dx.doi.org/10.1016/s0028-3932(01)00014-8.

Kircher, T. T. J., Brammer, M. J., Williams, S. C. R., & McGuire, P. K. (2000). Lexical retrieval during fluent speech production: An fMRI study. *NeuroReport, 11*(18), 4093–4096.

Kircher, T. T. J., Oh, T. M., Brammer, M. J., & McGuire, P. K. (2005). Neural correlates of syntax production in schizophrenia. *The British Journal of Psychiatry, 186*(3), 209–214. http://dx.doi.org/10.1192/bjp. 186.3.209.

Krekelberg, B., Boynton, G. M., & Wezel, R. J. A. v. (2006). Adaptation: From single cells to BOLD signals. *Trends in Neurosciences, 29*(5), 250–256.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press. Levelt's original book.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1–75.

Long, D. L., Baynes, K., & Prat, C. S. (2005). The propositional structure of discourse in the two cerebral hemispheres. *Brain and Language, 95*, 383–394.

Martin, R. C. (2003). Language processing: Functional organization and neuroanatomical basis. *Annual Review of Psychology, 54*, 55–89.

Meltzer, J. A., McArdle, J. J., Schafer, R. J., & Braun, A. R. (2010). Neural aspects of sentence comprehension: Syntactic complexity, reversibility, and reanalysis. *Cerebral Cortex, 20*(8), 1853–1864. http://dx.doi.org/10.1093/cercor/bhp249.

Menenti, L., Gierhan, S. M. E., Segaert, K., & Hagoort, P. (2011). Shared language. *Psychological Science, 22*(9), 1174–1182. http://dx.doi.org/10.1177/0956797611418347.

Menenti, L., Petersson, K. M., & Hagoort, P. (2012). From reference to sense: how the brain encodes meaning for speaking [original research]. *Frontiers in Psychology, 2*. http://dx.doi.org/10.3389/fpsyg.2011.00384.

Mitchell, R. L. C., & Crow, T. J. (2005). Right hemisphere language functions and schizophrenia: The forgotten hemisphere? *Brain, 128*, 963–978.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113. doi:citeulike-article-id:1006238.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience, 8*(12), 976–987.

Poser, B. A., Versluis, M. J., Hoogduin, J. M., & Norris, D. G. (2006). BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: Parallel-acquired inhomogeneity desensitized fMRI. *Magnetic Resonance in Medicin, 55*, 1227–1235.

Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences of the United States of America, 107*(32), 14425–14430. http://dx.doi.org/10.1073/pnas.1008662107.

Tremblay, P., & Gracco, V. L. (2006). Contribution of the frontal lobe to externally and internally specified verbal responses: fMRI evidence. *NeuroImage, 33*(3), 947–957. http://dx.doi.org/10.1016/j.neuroimage.2006.07.041.

Tremblay, P., & Gracco, V. L. (2010). On the selection of words and oral motor responses: Evidence of a response-independent fronto-parietal network. *Cortex, 46*(1), 15–28. http://dx.doi.org/10.1016/j.cortex.2009.03.003.

Tremblay, P., & Small, S. L. (2011a). From language comprehension to action understanding and back again. *Cerebral Cortex, 21*(5), 1166–1177. http://dx.doi.org/10.1093/cercor/bhq189.

Tremblay, P., & Small, S. L. (2011b). On the context-dependent nature of the contribution of the ventral premotor cortex to speech perception. *NeuroImage, 57*(4), 1561–1571. http://dx.doi.org/10.1016/j.neuroimage.2011.05.067.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage, 15*, 273–289.

Vigliocco, G., & Hartsuiker, R. J. (2002). The interplay of meaning, sound and syntax in sentence production. *Psychological Bulletin, 128*(3), 442–472. doi:10.1037//0033-2909.128.3.442.

Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. *Cognition, 75*(2), 105–143.

Weber, K., & Indefrey, P. (2009). Syntactic priming in German-English bilinguals during sentence comprehension. *NeuroImage, 46*(4), 1164–1172.

Wilson, S. M., Ogar, J. M., Laluz, V., Growdon, M., Jang, J., Glenn, S., et al. (2009). Automated MRI-based classification of primary progressive aphasia variants. *NeuroImage, 47*(4), 1558–1567. doi:10.1016/j.neuroimage.2009.05.085.

Wise, R. J. S., Scott, S. K., Blank, S. C., Mummery, C. J., Murphy, K., & Warburton, E. A. (2001). Separate neural subsystems within 'Wernicke's area'. *Brain, 124*(1), 83–95. http://dx.doi.org/10.1093/brain/124.1.83.

Zheng, Z. Z., Munhall, K. G., & Johnsrude, I. S. (2010). Functional overlap between regions involved in speech perception and in monitoring one's own voice during speech production. *Journal of Cognitive Neuroscience, 22*(8), 1770–1781.