Neural basis of linearization in speech production

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Neural Basis of Linearization in Speech Production

Zheng Ye[1,2]*, Boukje Habets[3]*, Bernadette M. Jansma[4],
and Thomas F. Münte[2]

Abstract

■ An initial stage of speech production is conceptual planning, where a speaker determines which information to convey first (the linearization problem). This fMRI study investigated the linearization process during the production of "before" and "after" sentences. In "after" sentences, a series of events is expressed in the order of event occurrence. In "before" sentences, however, the order of event mention is achieved by reversing the chronological order. We suggested that the linearization process may be supported by a neural network connecting the left middle temporal gyrus (MTG) with the medial superior frontal gyrus, left middle frontal gyrus, and left angular gyrus/inferior parietal gyrus. Within this network, regions were more activated and interregional interactions were strongly enhanced for producing "before" than "after" sentences. The left MTG was also functionally connected with the left orbital inferior frontal gyrus, contributing to the retrieval of necessary world knowledge and linguistic knowledge. Connectivity between these two regions was not different between conditions. ■

INTRODUCTION

When asked to describe a series of events, a speaker is able to start an utterance such as "after I ate dinner, I cleaned dishes" in just a few hundred milliseconds. According to psycholinguistic models of speech production, the speaker needs to pass through a number of processing stages in this brief time, including conceptual preparation (for a description of conceptual representation, see Mahon & Caramazza, 2009), followed by grammatical, morpho-phonological, and phonetic encoding (Caramazza, 1997; Levelt, 1989; Dell, 1986). Although all of these models feature access to these different kinds of information, they differ in their specification of the cognitive architecture, in particular, with regard to whether they allow bidirectional interaction between levels (e.g., Dell, 1986) or not. In broad terms, they all make similar assumptions on the general architecture of a speech production system, such as assuming separate conceptual and lexical subprocesses that are interconnected. In Levelt's (1989, 1999) model, for example, a message is initially generated through macro- and micro-planning during conceptualization. In macro-planning, the communicative intention is elaborated with a sequence of speech acts (i.e., actions one performs by speaking, such as informing, requesting, and apologizing). To perform a speech act, the speaker needs not only to determine which information to communicate but also to decide on the order in which the information will be conveyed (the linearization problem). He or she can mention events in the order of occurrence using the sentence-initial temporal connective "after" (e.g., "**After** I sit down, I open the mail") or in the reversed order with "before" (e.g., "**Before** I open the mail, I sit down"). The speech act is further specified in micro-planning, where the speaker selects the perspective of an utterance. For example, he or she can refer to a person with a pronoun (e.g., "she"), if this person has appeared in the preceding context, or with a noun phrase (e.g., "an actress"), if the person has never been mentioned in the ongoing conversation before. The output of macro- and micro-planning is a preverbal plan, which contains all information necessary for converting an idea into a meaningful utterance. The preverbal plan then enters into grammatical encoding, where appropriate words are retrieved from the mental lexicon and arranged in a syntactic frame. During later processing stages, words are specified with morphological, metrical, and segmental features (e.g., stress and pitch), and finally, transferred into codes which can be executed by the articulatory system.

Empirical support for the serial/cascading processing is based mainly on single noun production during picture naming. Behavioral studies suggested a sequential availability of semantics and phonology (Levelt et al., 1991; Schriefers, Meyer, & Levelt, 1990) and a left-to-right sequential phonological retrieval (Wheeldon & Levelt, 1995; Meyer, 1990, 1991). A general difficulty in speech production is to disentangle the different subprocesses experimentally. Electrophysiology (i.e., ERP) research has contributed to the insights of the time course of speaking, taking advantage of its exquisite temporal resolution. Findings from ERP studies using the N200 paradigm further supported the idea of cascading processing (Schiller, Bles, & Jansma,

[1]Capital Medical University, Beijing, China, [2]University of Lübeck, Germany, [3]University of Hamburg, Germany, [4]Maastricht University, The Netherlands
*Z. Y. and B. H. contributed equally to this work.

2003; Rodriguez-Fornells, Schmitt, Kutas, & Münte, 2001; Schmitt, Münte, & Kutas, 2000). Using lateralized readiness potentials, van Turennout, Hagoort, & Brown (1998) showed that syntactic encoding precedes phonological encoding. In a meta-analysis of behavioral, ERP, and fMRI studies, Indefrey and Levelt (2004) described which brain areas are active at a certain time during the planning of overt speech. Overall, the idea is that conceptual encoding takes place up to 150 msec after picture onset (Thorpe, Fize, & Marlot, 1996), followed by lexical and syntactic information retrieval (200–300 msec). Phonological retrieval starts around 300 msec poststimulus onset and the entire phonological plan is available around 450 msec after picture onset. At around 600 msec, an articulatory plan is ready for articulation preparation in motor cortex. The cascading processing assumption is still under debate as reflected by behavioral studies (Caramazza & Miozzo, 1997), neuropsychological case studies (Shapiro & Caramazza, 2003), and ERP studies (Abdel Rahman & Sommer, 2003). The suggested anatomical network, including medial frontal cortex, cingulate cortex, middle frontal gyrus (MFG), inferior frontal gyrus (IFG, Broca's area), superior temporal gyrus, middle temporal gyrus (MTG), and inferior temporal gyrus, continues to be revised and updated, based on recent evidences for example on conceptual versus category encoding (Mahon & Caramazza, 2009), or with regard to the role of Broca's area (Hagoort, 2005). Recently, using transcranial magnetic stimulation, it has been confirmed that phonological planning takes place in Broca's area, and that syllabication can be hampered when a virtual lesion was applied within Broca's around 300 msec after picture onset. This delay in naming effect did not occur when the magnetic pulse was applied elsewhere in the language network, or when the magnetic pulse was given earlier or later in time (Schumann, Schiller, Goebel, & Sack, 2009). In addition, using invasive intercranial local field potentials, Sahin, Pinker, Cash, Schomer, and Halgren (2009) confirmed the relevance of Broca's area and the cascading temporal processing in overt speech for semantics, syntax, and phonological processing.

Because of the difficulty in controlling the conceptual planning of complex utterances, previous neuroimaging studies concentrated on conceptual representation and lexical selection during single-word production (e.g., Crosson et al., 2001; Palmer et al., 2001; Etard et al., 2000; for reviews, see Indefrey & Levelt, 2004; Price, 1998), or syntactic encoding during simple sentence production (e.g., Haller, Radue, Erb, Grodd, & Kircher, 2005; Indefrey, Hellwig, Herzog, Seitz, & Hagoort, 2004; Indefrey et al., 2001). These studies revealed the medial frontal cortex, left IFG, premotor cortex, left anterior temporal area, posterior superior temporal gyrus, and inferior parietal lobule (IPL) on the word level, and the medial frontal cortex, ACC, left IFG, premotor cortex, left MTG, and left IPL on the sentence level. A first study on the neural mechanisms of conceptualization in speech production was carried out by Habets, Jansma, and Münte (2008). In that ERP study, Habets

et al. developed a paradigm which allows the assessment of the linearization process in macro-planning. In this paradigm, participants see a sequence of two objects, each associated with one action, and then a colored cue. Participants are told that the first action associated with the first object occurred first and the second action associated with the second object happened subsequently. They are asked to describe the sequence of actions with sentence-initial temporal connective "before" or "after," according to the color of the cue. For example, if they see a first object "book" and a second object "couch," followed by a yellow fixation cross, they should utter a German sentence like "**Bevor** ich sitze, lese ich [**Before** I sit (down on the couch), I read (a book); information in parentheses for clarification only]." If they see the same pair of objects followed by a red fixation cross, they should utter a German sentence like "**Nachdem** ich lese, sitze ich [**After** I read (a book), I sit (down on the couch)]." For "after" sentences, participants can arrange the two constituent clauses in the order of event occurrence. For "before" sentences, they have to reverse the chronological order to reach the correct linguistic order, which probably requires more processing resources and/or additional computations. Habets et al. found brain responses in the "before" and the "after" conditions diverged around 180 msec after the presentation of the cue. The "before" condition was less negative around 200 msec over the fronto-central scalp but more positive around 350 msec over the parieto-occipital scalp. Habets et al. suggested that the latter positivity could be an instance of the P300 or P3b, which reflects the greater demand for processing during the building of "before" sentences. Their observations on production are in line with existing findings on acquisition and comprehension. "Before" sentences have been shown to be more difficult to understand, in particular for young children (Natsopoulos & Abadzi, 1986; Trosberg, 1982; Clark, 1971), patients with Parkinson's disease (Natsopoulos et al., 1991), or patients with aphasia (Sasanuma & Kamio, 1976). Although "before" sentences can be understood as well as "after" sentences in healthy adults (Natsopoulos et al., 1991), they give rise to a sustained negative effect over the left frontal scalp in ERPs (Münte, Schiltz, & Kutas, 1998) and more activations in medial frontal cortex, MFG, MTG, hippocampus, and basal ganglia in a recent neuroimaging study of our laboratory (Ye, Kutas, St. George, Sereno, & Münte, submitted).

In the current study, we examined which brain areas support the linearization process in the production of "before" and "after" sentences using fMRI. The design of Habets et al. (2008) was adjusted for fMRI measurements. We used a standard univariate analysis to investigate regions more activated for producing "before" than "after" sentences followed by a functional connectivity analysis to define the brain networks engaged by each condition. The connectivity approach is based on the "beta series correlation" method proposed by Rissman, Gazzaley, & D'Esposito (2004), which has been successfully applied in studies of different cognitive domains (e.g.,

Camara, Rodriguez-Fornells, & Münte, 2008; Rissman, Gazzaley, & D'Esposito, 2008; Fiebach, Rissman, & D'Esposito, 2006). This method is implemented on the basis of a general linear model (GLM), using separate covariates to model hemodynamic responses of a particular stage (e.g., producing sentences or viewing objects) in each single trial, and then generating a series of parameter estimates (beta values) for that stage. If two regions are functionally connected within a network, their beta series should be strongly correlated. A neural network would be considered to support the linearization process, if the connectivity in this network was stronger for producing "before" than "after" sentences.

METHODS

All procedures had been cleared by the ethical review board of the University of Magdeburg.

Participants

Fifteen native German speakers (8 women, mean age = 23.8 years, age range = 21 to 29 years) participated in this study. They were right-handed and had normal or corrected-to-normal vision. None of them had a history of neurological or psychiatric disorder. All of them gave written informed consent before scanning.

Stimuli and Task

In each trial, participants saw a sequence of two objects and overtly described the actions associated with the objects using sentences starting with a temporal connective, "before" or "after." Forty objects were presented using black–white line drawings from the picture data base of the Max-Planck Institute for Psycholinguistics, Nijmegen, or the Snodgrass and Vanderwart (1980) collection. Pictures were resized into 33 × 33 mm with a resolution of 300 × 300 dpi, and were combined into pairs so that there is no semantic or phonological overlap between the German names of the two objects presented in a trial. Each object pair was presented twice, once per condition, with the presentation order reversed on the second presentation. In other words, there were 40 pairs of objects in each condition.

The two objects in a pair were sequentially presented. Each object was displayed for 500 msec, separated by an interstimulus interval of 500 msec. Then, a colored fixation cross (yellow/red) stayed on the screen for 4000 msec to allow for an overt response. The fixation color was counterbalanced across participants. The colored fixation cross was replaced by a 14-sec white fixation cross, indicating that the response time was over. Participants were instructed to utter the required sentence as quickly and accurately as possible. In order to optimally match the two types of utterances, participants were asked to use an identical syntactic structure in both conditions (except for the initial word) and to produce both constituent

clauses in the present tense. Please note that although German speakers most commonly would choose the past participle (*Nachdem ich gelesen habe, habe ich gebügelt*), utterances in the present tense also sound natural for German speakers and comprehenders. With the present tense we were able to minimize variability between answers and to minimize the length of the produced utterance. Lip movements were recorded with a camera to monitor whether participants used the correct sentence initial temporal connective (to pronounce *bevor*, the mouth is closed at the beginning; to pronounce *nachdem*, the mouth is slightly open). There were four functional runs, each lasting about 7 min. Participants received three practice runs before scanning to learn the corresponding verbs to the presented objects and to get familiar with the task.

Data Acquisition

Data were collected on a 3-T Siemens Trio system. Functional images were acquired using a T2*-weighted echo-planar imaging (EPI) sequence, with a 2000-msec time repetition (TR), a 30-msec time echo (TE), and an 80° flip angle, in four runs. Each functional image consisted of 30 axial slices, with 64 × 64 matrix, 220 × 220 mm field of view (FOV), 3.5-mm thickness, 0.35-mm gap, and 3.5 × 3.5 mm in-plane resolution. Structural images were acquired using a T1-weighted magnetization-prepared rapid-acquired gradient-echo (MP-RAGE) sequence, with 2500-msec TR, 1.68-msec TE, and 7° flip angle. The structural image consisted of 192 slices, with 256 × 256 matrix, 256 × 256 mm FOV, 1-mm thickness, no gap, and 1 × 1 mm in-plane resolution.

Data Analysis

Data were analyzed with SPM5 (www.fil.ion.ucl.ac.uk/spm). The first four volumes of each run were discarded, owing to equilibration effects. Functional images were first phase-shifted with reference to the middle slice to correct differences in slice acquisition time. They were then realigned with a least squares approach and a rigid-body spatial transformation to remove artifacts. Estimated movement parameters (six per image: x, y, z, pitch, roll, and yaw) were included in statistical analysis to minimize signal-correlated motion effects. Realigned images were normalized to the EPI-derived MNI template (ICBM 152, Montreal Neurological Institute) and resampled to $2 \times 2 \times 2$ mm³ voxel size. Normalized images were finally smoothed with a Gaussian kernel of 8-mm full-width half-maximum and filtered with a high-pass filter of 128 sec. We conducted two statistical analyses, namely, a standard univariate analysis and a functional connectivity analysis.

Standard Univariate Analysis

The standard univariate analysis was performed to examine which regions were more activated for producing

"before" than "after" sentences, and vice versa. Hemodynamic responses were modeled on the basis of a GLM with a canonical hemodynamic response function. Three types of events were specified in the model, including the pictures of two objects, the "after" sentence, and the "before" sentence. The picture event was time-locked to the onset of the first object picture. The sentence event was time-locked to the onset of the colored fixation cross. Classical parameter estimation was applied with a one-lag autoregressive model to whiten temporal noise in order to correct the probability (p) of a false-positive voxel on the subject level. With this subject-level correction, there is no need for correction on the group level (Smith, Singh, & Balsters, 2007; Bianciardi, Cerasa, Patria, & Hagberg, 2004). Thus, we reported p values in following random-effect analysis without further correction. For each participant, a contrast map was calculated for the "before" versus "after" condition. Contrast maps were entered into a one-sample t test on the group level (multisubject random effect). Activations were considered significant at $p < .01$ with a minimum cluster of 20 voxels.

Functional Connectivity Analysis

The connectivity analysis was performed to examine how different areas worked together during the production of "before" and "after" sentences. This approach is based on the hypothesis that if two regions interact within a network, their activity patterns should be strongly correlated (Rissman et al., 2004). This analysis was implemented on the basis of a GLM by using separate covariates to model hemodynamic responses of a particular stage (e.g., viewing picture or producing sentence) in each single trial. For each condition, beta values (estimated parameter of activity strength) were extracted to form a set of condition-specific beta series. The left MTG ROI was used as the seed because it was the most consistently activated area in the standard analysis. Beta series of the seed were averaged across voxels within the critical region and correlated with beta series of every other voxel in the whole brain. For each participant, a map of correlation coefficients was calculated for each condition. All correlation maps were normalized by using an arc-hyperbolic tangent transform for further statistical inference. To show connectivity patterns in both conditions, we first conducted one-sample t tests separately for the "before" and the "after" conditions on the group level. Resulting maps were considered at $p < .05$ (family-wise error correction, FWE-corrected) with a minimum cluster size of 100 voxels. An FWE is a false positive anywhere in the resulting map. A value of .05 implies that if one repeats an experiment many times, only 1 in 20 resulting maps contains a false positive somewhere in the image. Thus, FWE-correction is a very conservative threshold. The resulting maps showed regions significantly correlated with the left MTG in beta series and were inferred to be functionally connected with the left MTG in each condition. We then conducted a flexible factorial test to show the con-

nectivity stronger in the "before" versus "after" condition. The flexible factorial test in SPM5 allows examining the difference between two conditions on the group level without computing contrast maps on the subject level. The resulting map was considered at $p < .05$ with a minimum cluster size of 20 voxels.

RESULTS

Response Errors

Error rates were 1.2% ($SE = 0.6\%$) in the "after" condition and 1.7% ($SE = 0.5\%$) in the "before" condition ($t = -0.72$, $df = 14$, $p = .49$, 2-tailed), indicating that participants performed equally well in both conditions.

fMRI Results

The standard univariate analysis revealed regions more activated for producing "before" than "after" sentences (see Figure 1 and Table 1), including the ACC, MFG, left orbital inferior frontal gyrus/superior temporal pole (IFGorb/STP), MTG, and angular gyrus/inferior parietal lobule (AG/IPL). None of these regions showed up in the "after" versus "before" comparison, which revealed activity in the right supplementary motor area (MNI: 8 −22 60, $t = 4.37$, $p < .001$, 139 voxels), right postcentral gyrus (MNI: 12 −32 64, $t = 4.12$, $p = .001$, 61 voxels), and right insular cortex (MNI: 38 −10 −2, $t = 3.21$, $p = .003$, 62 voxels).

The functional connectivity analysis revealed regions interacting with the seed region during the production of "before" and "after" sentences (see Figure 2 and Table 2). The left MTG was functionally connected with the mSFG/ACC, left MFG, left IFGorb, and left AG/IPL. Figure 2 also shows interregional interactions stronger in the "before" versus "after" condition. The MTG connections were significantly enhanced in the "before" condition with the mSFG ($t = 2.77$, $df = 14$, $p = .005$, 578 voxels), the left MFG ($t = 2.18$, $df = 14$, $p = .019$, 52 voxels), and the left AG/IPL ($t = 2.66$, $df = 14$, $p = .006$, 252 voxels), but not with the left IFGorb (neither in the "after" vs. "before" comparison).

DISCUSSION

In the present study, we delineated the brain areas underlying the linearization process during speech production, including the mSFG/ACC, left MFG, left IFGorb, left MTG, and left AG/IPL. These regions were more activated for producing "before" sentences, in which the constituent clauses are ordered differently from the chronological order of events, than for producing "after" sentences, in which the order of event mention is identical to the order of events in the real world. Moreover, these regions worked together as a left-lateralized neural network as revealed by functional connectivity analysis. Within this
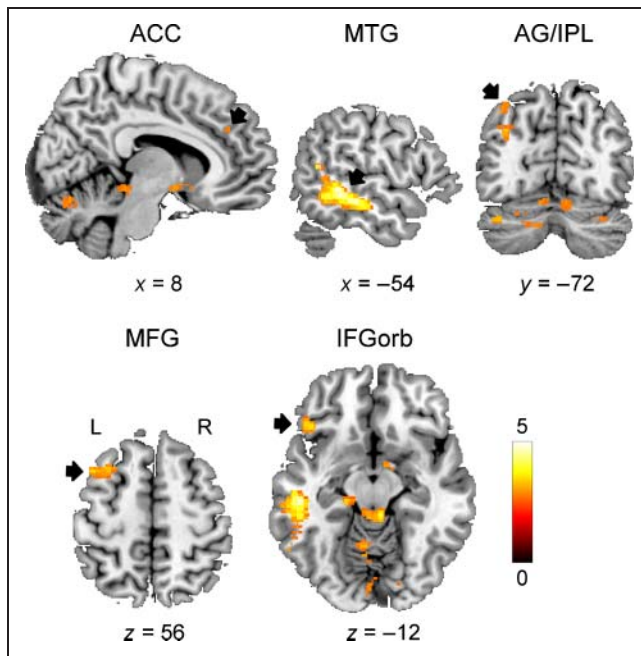
Figure 1. Regions more activated for producing "before" than "after" sentences. Slices were selected to show the anterior cingulate cortex (ACC), middle frontal gyrus (MFG), orbital inferior frontal gyrus (IFGorb), left middle temporal gyrus (MTG), and angular gyrus/inferior parietal lobule (AG/IPL). Arrows point to activated regions. Color scale indicates *t* values. Coordinates in MNI; L = left; R = right.

network, the left MTG's connectivity with the mSFG, left MFG, and left AG/IPL was enhanced in the "before" versus "after" condition. However, the connection between the left MTG and the left IFGorb was not different between conditions. This connectivity pattern suggests that different nodes of this network may be responsible for different functions.

Brain areas such as the left MTG, mSFG, left MFG, and left IFG have consistently been observed in picture naming and word generation tasks (Indefrey & Levelt, 2004). They are assumed to be crucial for the core processes of language production with different regions supporting different functions. The left MTG is associated with the conceptually driven retrieval and selection of lexical items and the left IFG is linked to syllabification and metrical encoding (also see Indefrey & Levelt, 2000). Previous studies also suggested that the left AG plays a role in semantic processes during word reading and picture naming (Mechelli, Josephs, Ralph, McClelland, & Price, 2007), as cortical stimulation of the angular and supramarginal gyri has been shown to elicit delays in verbal responses, misnaming, and difficulties in speech production (Van Buren, Fedio, & Frederick, 1978). Recent studies further observed the involvement of medial frontal cortex in syntactic encoding (Haller et al., 2005) and lexical retrieval (Crosson et al., 2001). For example, participants showed more medial frontal

Table 1. Regions More Activated for Producing "Before" than "After" Sentences

Region	BA	H	x	y	z	t	No. of Voxels
Anterior cingulate cortex	32	L	12	38	30	3.30	21
Middle frontal gyrus	9	L	−40	14	56	3.63	220
		R	32	24	44	3.57	84
Inferior frontal gyrus (pars triangularis)	45	R	50	26	24	3.43	105
Inferior frontal gyrus (pars orbitalis)	47	L	−34	28	−22	3.65	212
Superior temporal pole	38	L	−44	24	−16	3.94	
Middle temporal gyrus	21/20	L	−54	−40	−6	5.18	895
Inferior temporal gyrus	20	L	−46	−2	−34	3.13	77
Precuneus/superior parietal lobule	7/5	L	−10	−52	76	3.49	65
		R	22	−62	70	3.38	30
Angular gyrus/inferior parietal gyrus/	39/7	L	−38	−72	44	3.13	46
Middle occipital gyrus	19	L	−34	−72	32	3.63	368
Insula		L	−26	30	−2	3.88	53
		R	34	32	0	3.14	31
Hippocampus		L	−20	−32	10	5.15	532
		R	22	−32	12	4.55	250
Cerebellum		L	−40	−68	−30	4.20	335
		R	32	−68	−32	3.48	107

BA = approximate Brodmann's area; H = hemisphere; *t* = statistic values of peaks; L = left; R = right; coordinates are in MNI.

Figure 2. Connectivity patterns in the "before" and "after" conditions and differences in connectivity between conditions. Slices were selected to show the medial superior frontal cortex/anterior cingulate cortex (mSFG/ACC), middle frontal gyrus (MFG), orbital inferior frontal gyrus (IFGorb), and angular gyrus/inferior parietal lobule (AG/IPL). Arrows point to connected regions. Color scale indicates t values. Coordinates in MNI; L = left; R = right.
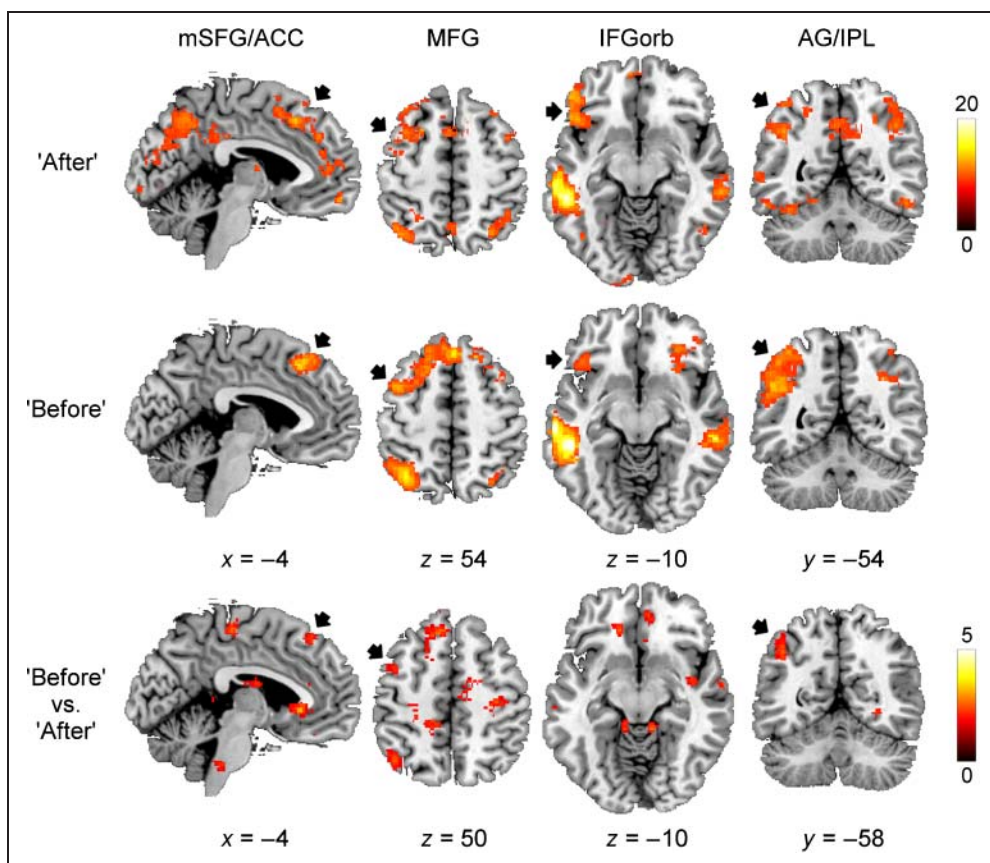
activation when generating a sentence (e.g., "the child throws the ball") from an unordered list of words (e.g., "throw ball child") than simply reading the word list or reading the sentence (Haller et al., 2005). The fact that medial frontal cortex has been observed both for ordering words in a syntactic frame (Haller et al.'s study) and for ordering concepts to reach a preverbal plan (the current study) suggests that this area may play a role in sequence processing. This assumption is in line with previous findings that medial frontal cortex is involved in learning and processing sequences of linguistic and/or nonlinguistic stimuli (e.g., Bahlmann, Schubotz, Mueller, Koester, & Friederici, 2009; Forkstam, Hagoort, Fernandez, Ingvar, & Petersson, 2006; Lehéricy et al., 2006). Moreover, both ACC and the IPL have been observed for overt word-stem completion (Palmer et al., 2001) and picture naming (de Zubicaray, McMahon, Eastburn, & Wilson, 2002).

In the perception and attention domain, the mSFG/ACC and IPL are usually called upon for maintaining and updating task demands, for monitoring and adjusting cognitive processes, and for allocating attentional resources (Liston, Matalon, Hare, Davidson, & Casey, 2006; Botvinick, Braver, Barch, Carter, & Cohen, 2001; Gitelman et al., 1999; Peterson et al., 1999; Carter et al., 1998). In the current study, the greater activations and the stronger interregional interactions for "before" sentences may reflect that more processing resources and/or additional

computations are employed to achieve the order of event mention by reversing the order of event occurrence. The left MFG node of this network may serve as a workspace where information contributed by other regions can be manipulated and integrated, because this structure has been identified with working memory processes (Tsukiura et al., 2001; McCarthy et al., 1994, 1996). The network identified in the current study may underlie the late positive effect elicited by "before" utterances as observed in Habets et al. (2008).

Different from other regions in the network, the left IFGorb did not show stronger connectivity with the left MTG in the "before" condition. The left IFGorb is assumed to support controlled access to stored conceptual and episodic representations in word production and comprehension (Badre & Wagner, 2007; Badre, Poldrack, Paré-Blagoev, Insler, & Wagner, 2005). Previous studies found that this region is anatomically connected with the anterior temporal area (Anwander, Tittgemeyer, von Cramon, Friederici, & Knösche, 2007) and functionally coupled with the left MTG in the processing of regularly and irregularly inflected words (Stamatakis, Marslen-Wilson, Tyler, & Fletcher, 2005). The similar connectivity patterns in the two conditions suggest that the memory processes were engaged to a similar extent for "before" and "after" sentences. Alternatively, the memory processes may be employed to maintain a pair of concepts (i.e., object pictures) in working

Table 2. Regions Functionally Connected with the Left MTG during the Production of "Before" and "After" Sentences

Region	BA	H	After				Before			
			x	y	z	t	x	y	z	t
Medial superior frontal gyrus/supplementary motor area	8	L	−2	22	46	13.97	−4	24	48	17.11
Anterior cingulate cortex	32	L	−6	42	20	10.96				
Middle frontal gyrus	46	L	−30	24	34	17.67	−42	6	48	15.53
		R	48	34	32	10.83				
Inferior frontal gyrus (pars orbitalis)	47	L	−48	40	−12	15.69	−42	26	−8	11.35
		R					28	34	−10	11.99
Middle temporal gyrus	21/20	L	−54	−38	−2	29.00	−52	−38	−8	31.05
		R	64	−34	0	12.32	58	−32	−10	17.22
Inferior temporal gyrus	20	L	56	−18	−24	13.62				
		R	50	−20	−24	14.37				
Precuneus	7	R	4	−72	38	13.07				
Angular gyrus/Inferior parietal lobule	39/40	L	−44	−50	36	15.12	−48	−50	34	15.52
		R	30	−66	54	11.73	42	−50	34	12.09
Calcarine	17	R	4	−82	4	12.73	18	−76	12	12.86
Caudate nucleus		L	−10	16	4	13.89				
Thalamus		L	−16	−10	14	14.34				
Cerebellum		R	34	−64	−32	12.79	36	−66	−32	15.25

BA = approximate Brodmann's area; H = hemisphere; t = statistic values of peaks; L = left; R = right; coordinates are in MNI.

memory prior to the conceptual planning in the current paradigm.

Conclusion

In conclusion, we found a neural network in support of conceptual planning during speech production comprising the left MTG, mSFG, left MFG, and left AG/IPL. Within this network, regions were more activated and interregional interactions were strongly enhanced for producing "before" than "after" sentences. The activity patterns and the connectivity patterns consistently suggested that more processing resources and/or additional computations may be employed for "before" sentences, to reach the linguistic order of constituent clauses by reversing the chronological order of events. The left MTG was also functionally connected with the left IFGorb, contributing to the retrieval of world knowledge and linguistic knowledge necessary for language production. However, please note that these areas are involved in, but probably not specific for, the linearization process. Areas such as the mSFG mainly play a role in attentional control and/or sequence processing. Areas such as the left AG/IPL only showed a subtle effect in activity. Further studies are necessary before reaching a conclusion.

REFERENCES

Abdel Rahman, R., & Sommer, W. (2003). Does phonological encoding in speech production always follow the retrieval of semantic knowledge? Electrophysiological evidence for parallel processing. *Cognitive Brain Research, 16,* 372–382.

Anwander, A., Tittgemeyer, M., von Cramon, D. Y., Friederici, A. D., & Knösche, T. R. (2007). Connectivity-based parcellation of Broca's area. *Cerebral Cortex, 17,* 816–825.

Badre, D., Poldrack, R. A., Paré-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron, 47,* 907–918.

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia, 45,* 2883–2901.

Bahlmann, J., Schubotz, R. I., Mueller, J. L., Koester, D., & Friederici, A. D. (2009). Neural circuits of hierarchical

visuo-spatial sequence processing. *Brain Research, 1298,* 161–170.

Bianciardi, M., Cerasa, A., Patria, F., & Hagberg, G. E. (2004). Evaluation of mixed effects in event-related fMRI studies: Impact of first-level design and filtering. *Neuroimage, 22,* 1351–1370.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108,* 624–652.

Camara, E., Rodriguez-Fornells, A., & Münte, T. F. (2008). Functional connectivity of reward processing in the brain. *Frontiers in Human Neuroscience, 2,* 19.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology, 14,* 177–208.

Caramazza, A., & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: Evidence from the "tip-of-the-tongue" phenomenon. *Cognition, 64,* 309–343.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the on-line monitoring of performance. *Science, 280,* 747–749.

Clark, E. V. (1971). On the acquisition of the meaning of "before" and "after". *Journal of Verbal Learning and Verbal Behavior, 10,* 266–275.

Crosson, B., Sadek, J. R., Maron, L., Gökçay, D., Mohr, C. M., Auerbach, E. J., et al. (2001). Relative shift in activity from medial to lateral frontal cortex during internally versus externally guided word generation. *Journal of Cognitive Neuroscience, 13,* 272–283.

de Zubicaray, G. I., McMahon, K. L., Eastburn, M. M., & Wilson, S. J. (2002). Orthographic/phonological facilitation of naming responses in the picture–word task: An event-related fMRI study using overt vocal responding. *Neuroimage, 16,* 1084–1093.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93,* 283–321.

Etard, O., Mellet, E., Papathanassiou, D., Benali, K., Houde, O., Mazoyer, B., et al. (2000). Picture naming without Broca's and Wernicke's area. *NeuroReport, 11,* 617–622.

Fiebach, C. J., Rissman, J., & D'Esposito, M. (2006). Modulation of inferotemporal cortex activation during verbal working memory maintenance. *Neuron, 21,* 251–261.

Forkstam, C., Hagoort, P., Fernandez, G., Ingvar, M., & Petersson, K. M. (2006). Neural correlates of artificial syntactic structure classification. *Neuroimage, 32,* 956–967.

Gitelman, D. R., Nobre, A. C., Parrish, T. B., LaBar, K. S., Kim, Y.-H., Meyer, J. R., et al. (1999). A large-scale distributed network for covert spatial attention: Further anatomical delineation based on stringent behavioural and cognitive controls. *Brain, 122,* 1093–1106.

Habets, B., Jansma, B. M., & Münte, T. F. (2008). Neurophysiological correlates of linearization in language production. *BMC Neuroscience, 9,* 77.

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

Haller, S., Radue, E. W., Erb, M., Grodd, W., & Kircher, T. (2005). Overt sentence production in event-related fMRI. *Neuropsychologia, 43,* 807–814.

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.

Indefrey, P., Hellwig, F., Herzog, H., Seitz, R. J., & Hagoort, P. (2004). Neural responses to the production and comprehension of syntax in identical utterances. *Brain and Language, 89,* 312–319.

Indefrey, P., & Levelt, W. J. M. (2000). The neural correlates of language production. In M. Gazzaniga (Ed.), *The new cognitive neurosciences*. Cambridge, MA: MIT Press.

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

Lehéricy, S., Bardinet, E., Tremblay, L., Van de Moortele, P.-F., Pochon, J.-B., Dormont, D., et al. (2006). Motor control in basal ganglia circuits using fMRI and brain atlas approaches. *Cerebral Cortex, 16,* 149–161.

Levelt, W. J. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. (1999). Models of word production. *Trends in Cognitive Sciences, 3,* 223–232.

Levelt, W. J. M., Schriefer, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review, 98,* 122–142.

Liston, C., Matalon, S., Hare, T. A., Davidson, M. C., & Casey, B. J. (2006). Anterior cingulated and posterior parietal cortices are sensitive to dissociable forms of conflict in a task-switching paradigm. *Neuron, 50,* 643–653.

Mahon, B. Z., & Caramazza, A. (2009). Concepts and categories: A cognitive neuropsychological perspective. *Annual Review of Psychology, 60,* 27–51.

McCarthy, G., Blamire, A. M., Puce, A., Nobre, A. C., Bloch, G., Hyder, F., et al. (1994). Functional magnetic resonance imaging of human prefrontal cortex activation during a spatial working memory task. *Proceedings of the National Academy of Sciences, U.S.A., 91,* 8690–8694.

McCarthy, G., Puce, A., Constable, R. T., Krystal, J. H., Gore, J. C., & Goldman-Rakic, P. (1996). Activation of human prefrontal cortex during spatial and nonspatial working memory tasks measured by functional MRI. *Cerebral Cortex, 6,* 600–611.

Mechelli, A., Josephs, O., Ralph, M. A. L., McClelland, J. L., & Price, C. J. (2007). Dissociating stimulus-driven semantic and phonological effect during reading and naming. *Human Brain Mapping, 28,* 205–217.

Meyer, A. S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language, 29,* 524–545.

Meyer, A. S. (1991). The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language, 30,* 69.

Münte, T. F., Schiltz, K., & Kutas, M. (1998). When temporal terms belie conceptual order. *Nature, 395,* 71–73.

Natsopoulos, D., & Abadzi, H. (1986). Understanding linguistic time sequence and simultaneity: A literature review and some new data. *Journal of Psycholinguistic Research, 15,* 243–273.

Natsopoulos, D., Mentenopoulos, G., Bostantzopoulou, S., Katsarou, Z., Grouios, G., & Logothetis, J. (1991). Understanding of relational time terms before and after in Parkinsonian patients. *Brain and Language, 40,* 444–458.

Palmer, E. D., Rosen, H. J., Ojemann, J. G., Buckner, R. L., Kelley, W. M., & Petersen, S. E. (2001). An event-related fMRI study of overt and covert word stem completion. *Neuroimage, 14,* 182–193.

Peterson, B. S., Skudlarski, P., Gatenby, J. C., Zhang, H., Anderson, A. W., & Gore, J. C. (1999). An fMRI study of Stroop word–color interference: Evidence for cingulate subregions subserving multiple distributed attentional systems. *Biological Psychology, 45,* 1237–1258.

Price, C. J. (1998). The functional anatomy of word comprehension and production. *Trends in Cognitive Sciences, 2,* 281–288.

Rissman, J., Gazzaley, A., & D'Esposito, M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. *Neuroimage, 23,* 752–763.

Rissman, J., Gazzaley, A., & D'Esposito, M. (2008). Dynamic adjustments in prefrontal, hippocampal, and inferior temporal interactions with increasing visual working memory load. *Cerebral Cortex, 18,* 1618–1629.

Rodriguez-Fornells, A., Schmitt, B. M., Kutas, M., & Münte, T. F. (2001). Electrophysiological estimates of the time course of semantic and phonological encoding during listening and naming. *Neuropsychologia, 40,* 778–787.

Sahin, N., Pinker, S., Cash, S. S., Schomer, D., & Halgren, E. (2009). Sequential processing of lexical, grammatical, and phonological information within Broca's area. *Science, 326,* 445–449.

Sasanuma, S., & Kamio, A. (1976). Aphasic's comprehension of sentences expressing temporal order of events. *Brain and Language, 3,* 495–506.

Schiller, N. O., Bles, M., & Jansma, B. M. (2003). Tracking the time course of phonological encoding in speech production: An event-related brain potential study. *Cognitive Brain Research, 17,* 819–831.

Schmitt, B. M., Münte, T. F., & Kutas, M. (2000). Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiology, 37,* 473–484.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language, 29,* 86–102.

Schumann, T., Schiller, N. O., Goebel, R., & Sack, A. T. (2009). The temporal characteristics of functional activation in Broca's area during overt picture naming. *Cortex, 45,* 1111–1116.

Shapiro, K., & Caramazza, A. (2003). Looming a loom: Evidence for independent access to grammatical and phonological properties in verb retrieval. *Journal of Neurolinguistics, 16,* 85–111.

Smith, A. T., Singh, K. D., & Balsters, J. H. (2007). A comment on the severity of the effects of non-white noise in fMRI time-series. *Neuroimage, 36,* 282–288.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experiment Psychology: Human Learning and Memory, 6,* 174–215.

Stamatakis, E. A., Marslen-Wilson, W. D., Tyler, L. K., & Fletcher, P. C. (2005). Cingulate control of fronto-temporal integration reflects linguistic demands: A three-way interaction in functional connectivity. *Neuroimage, 28,* 115–121.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381,* 520–522.

Trosberg, A. (1982). Children's comprehension of "before" and "after" reinvestigated. *Journal of Child Language, 9,* 381–402.

Tsukiura, T., Fujii, T., Takahashi, T., Xiao, R., Inase, M., Iijima, T., et al. (2001). Neuroanatomical discrimination between manipulating and maintaining processes involved in verbal working memory: A functional MRI study. *Cognitive Brain Research, 11,* 13–21.

Van Buren, J. M., Fedio, P., & Frederick, G. C. (1978). Mechanism and localization of speech in the parietotemporal cortex. *Neurosurgery, 2,* 233–239.

Van Turennout, M., Hagoort, P., & Brown, C. M. (1998). Brain activity during speaking: From syntax to phonology in 40 milliseconds. *Science, 280,* 572–574.

Wheeldon, L. R., & Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language, 34,* 311–334.

Ye, Z., Kutas, M., St. George, M., Sereno, M., & Münte, T. F. (submitted). Rearranging the world: Neural networks supporting the processing of temporal connectives.