Speaking of Which: Dissecting the Neurocognitive Network of Language Production in Picture Naming

Teresa Schuhmann1,2, Niels O. Schiller1,3,4, Rainer Goebel1,2 and Alexander Thomas Sack1,2

1Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, 6200 MD Maastricht, The Netherlands, 2Maastricht Brain Imaging Center, 6200 MD Maastricht, The Netherlands, 3Leiden Institute for Brain and Cognition, 2300 RA Leiden, The Netherlands and 4Netherlands Institute for Advanced Study (NIAS) in the Humanities and Social Sciences, 2242 PR Wassenaar, The Netherlands

Address correspondence to Teresa Schuhmann, Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands. Email: t.schuhmann@maastrichtuniversity.nl.

The noninvasive methods of cognitive neuroscience offer new possibilities to study language. We used neuronavigated multisite transcranial magnetic stimulation (TMS) to determine the functional relevance of 1) the posterior part of left superior temporal gyrus (Wernicke’s area), 2) a midportion of Broca’s area (slightly posterior/superior to apex of vertical ascending ramus), and 3) the midsection of the left middle temporal gyrus (MTG), during overt picture naming. Our chronometric TMS design enabled us to chart the time points at which neural activity in each of these regions functionally contributes to overt speech production. Our findings demonstrate that the midsection of left MTG becomes functionally relevant at 225 ms after picture onset, followed by Broca’s area at 300 ms and Wernicke’s area at 400 ms. Interestingly, during this late time window, the left MTG shows a second peak of functional relevance. Each area thus contributed during the speech production process at different stages, suggesting distinct underlying functional roles within this complex multicomponential skill. These findings are discussed and framed in the context of psycholinguistic models of speech production according to which successful speaking relies on intact, spatiotemporally specific feed forward and recurrent feedback loops within a left-hemispheric fronto-temporal brain connectivity network.

Keywords: Broca’s area, language production, picture naming, transcranial magnetic stimulation, Wernicke’s area

Introduction

Speaking is one of the most complex human skills. A seemingly simple task like naming an object requires the coordination of a series of processes, such as the selection of meanings, the retrieval of words, syntactic and phonological encoding, and articulation. The time course of these processes is described in models of speech production based on speech error data (Dell et al. 1997) and/or chronometric behavioral studies (Levelt et al. 1999).

More recently, imaging studies aimed at identifying which brain areas underlie the process of speaking. Based on a meta-analysis of 82 neuroimaging studies, Indefrey and Levelt (2004) suggested a detailed description of both the location and time course of cerebral activations during speech production. According to their neurocognitive model, the midsection of the left middle temporal gyrus (MTG) is functionally relevant first for lexical retrieval, followed by Wernicke’s area for phonological code retrieval, and finally Broca’s area for syllabification. The model, although influential in psycholinguistics, is still debated in the field of cognitive neuroscience, and so far no empirical brain research study has directly tested these spatio-temporal network predictions within controlled experimental conditions.

An elegant and most direct methodological approach to test these concrete predictions on a neuronal level would require controlled manipulation of local brain activity at specific time points during processing, with a quantifiable impact on speech production. Transcranial magnetic stimulation (TMS) enables such a manipulation of brain activity and is now a well-established tool for inducing transient disruptions of neural activity (virtual lesions) noninvasively in human volunteers. By transiently disrupting activity in the stimulated brain area and revealing a subsequent inability to perform a particular behavior, TMS can thus be regarded as a unique research tool for the investigation of causal structure-function relationships (Sack 2006).

In the current study, we applied chronometric TMS “online,” that is, while participants were performing a behaviorally controlled picture naming task, over 1) the posterior part of the left superior temporal gyrus (STG), which is part of Wernicke’s area, 2) a midportion of Broca’s area, located in the left, posterior inferior frontal gyrus (IFG), and 3) the midsection of the left MTG. Since TMS was applied to each of these magnetic resonance imaging (MRI)-identified brain regions at various time points between picture onset and overt speech production, at 150, 225, 300, 400, and 525 ms after picture presentation, we were able to address 2 independent research questions under controlled experimental conditions, namely unraveling whether intact neural activity in each of these brain areas is causally relevant for successful picture naming and at which precise points in time this neural activity in each area functionally contributes to the process of overt speech production.

Materials and Methods

Participants

Twelve healthy volunteers (5 men; mean age 23.2 years, standard deviation [SD] 2.08, ranging from 20 to 26 years) participated in the study. All participants were native Dutch speakers, had normal or corrected-to-normal vision, and had no history of neurological or psychiatric disorders. They received medical approval for participation and gave their written informed consent after being introduced to the procedure. The study was approved by the local Medical Ethical Committee.

Overall Study Design

Participants were tested in 4 separate sessions on 4 separate days. In the first session, we obtained anatomical brain measurements of all participants using MRI. We performed a surface reconstruction to
recover the spatial structure of the cortical sheet based on the white matter (WM) - gray matter (GM) boundary. We then identified 3 prominent language-related areas in the left hemisphere, namely IFG (Broca’s area), posterior STG (Wernicke’s area), and the midsection of the left MTG, on the basis of each individual brain gyration. In the following 3 sessions, we used frameless stereotaxy for MRI-guided TMS neuronavigation to target these regions with TMS and applied triple-pulse chronometric TMS during the execution of a behaviorally controlled picture naming task in order to study the influence of a controlled neural activity disruption on picture naming latencies. The order of stimulation site was randomized. This study design and methodological approach enabled us to first define the target brain area based on the individual anatomical data and to subsequently neuro-navigate the TMS coil to the anatomically defined stimulation sites in each participant. The MRI-guided TMS neuronavigation was monitored online throughout the experiment, allowing for a precise determination of the actual stimulation site during task execution.

**Stimuli, Paradigm, and Procedure**
A set of 10 simple white-on-black line drawings was used as target pictures. All items corresponded to monomorphemic monosyllabic Dutch nouns. They were taken from the picture database of the Max Planck Institute for Psycholinguistics in Nijmegen, The Netherlands. All picture names had a length between 3 and 5 segments (phonemes). Each picture had a mean frequency of occurrence between 10 and 73 per million as determined by CELEX (Baayen, Piepenbrock, and Gulikers 1995), that is, all target picture names were of moderate frequency. The drawings were presented on a computer screen in front of the participant. The stimuli subtended a visual angle of 2.82° × 4.57° and were displayed in the center of the monitor on a black background. Each trial consisted of a fixation cross presented between 5900 and 7900 ms, followed by a black screen for 100 ms. Thereafter, one of the pictures was presented for 750 ms. Participants were instructed to name the presented picture as quickly as possible by responding into a microphone. After a jittered delay between 6 and 8 s, the next trial began (see Fig. 1). We ruled out repetition priming by using only standard pictures that showed no repetition-related (implicit) learning effects (baseline plateau), as validated in psychophysical pilot measurements prior to the TMS study.

**Response Time Analysis**
The entire experiment was recorded with a microphone positioned on the table in front of the participant. Acoustic information was digitized with the digital audio editing software GoldWave v 5.17 (GoldWave, Newfoundland, Canada) with a sampling rate of 44 kHz. Prior to determining the speech onset, the acoustic signal was filtered to reduce noise. The latency of the verbal responses (reaction time) was measured on the screen using speech wave envelopes (see Fig. 2).

**MRI Measurements**
A high-resolution anatomical image was obtained from each participant in a 3-T magnetic resonance scanner (Siemens Allegra MR Tomograph; Siemens AG, Erlangen, Germany) at the Faculty of Psychology and Neuroscience, Maastricht University, The Netherlands. The T1-weighted data set was acquired with the help of a magnetization prepared rapid acquisition gradient echo sequence or a T1-weighted structural scan with an isotropic resolution of 1 mm using a modified driven equilibrium Fourier transform sequence with optimized contrast for GM and WM and imaging parameters.

**Cortical Surface Reconstruction**
Data were analyzed using the BrainVoyager QS 1.8 software package (Brain Innovation, Maastricht, The Netherlands). The high-resolution anatomical recordings were used for surface reconstruction of the left hemisphere of each participant (Kriegeskorte and Goebel 2001). The surface reconstruction was performed in order to recover the exact spatial structure of the cortical sheet and to improve the visualization of the anatomical gyification. The WM-GM boundary was segmented with a region growing method preceded by inhomogeneity correction of signal intensity across space. The borders of the resulting segmented subvolumes were tessellated to produce a surface reconstruction of the left hemisphere.

**Coregistration of Stereotaxic and MRI Data**
For a precise positioning of the stimulation coil, we made use of the BrainVoyager TMS Neuronavigator (Brain Innovation, Maastricht, The Netherlands). This neuronavigator system consists of several miniature ultrasound transmitters that are attached to a participant’s head, as well as the TMS coil. The ultrasound markers continuously transmit ultrasonic pulses to a receiving sensor device. The measurement of the relative spatial position of these transmitters in 3D space is based on travel time of the transmitted ultrasonic pulses to 3 microphones of the receiving sensor device. Local spatial coordinate systems are created by linking the relative raw spatial position of the ultrasound senders to a set of fixed additional landmarks on the participant’s head. The specification of these fixed landmarks is
achieved via a digitizing pen that also hosts 2 transmitting ultrasound markers in order to measure its relative position in 3D space. The 3 anatomical landmarks we used to define the local coordinate system were the nasion and the 2 incisure intertragicae. The neuronavigation system then provides topographic information of the head ultrasound transmitters relative to a participant-based coordinate frame. Similarly, the TMS coil also hosts a set of ultrasound transmitters whose relative spatial positions are linked to fixed landmarks specified on the coil in order to calculate another local coordinate system. After having defined the local spatial coordinate system for the participant’s head and the TMS coil in real 3D space, these coordinate systems have to be coregistered with the coordinate system of the MR space. For TMS-MRI coregistration, the same landmarks digitized on the participant’s head are specified on the head reconstruction of the anatomical data from the MR sequence. After the landmarks specified on the real head have been coregistered with those on the reconstructed head, events occurring around the head of the participant in real space are registered online and visualized in real time at correct positions relative to the participant’s anatomical reconstruction of the brain. The TMS coil can now be neuronavigated to a specific anatomical area of each participant. TMS neuronavigation was based on data in AC-PC space (rotating the cerebrum into the anterior commissure-posterior commissure plane) in order to avoid any additional transformations that could distort the correspondence between MRI and stereotactic points.

**TMS Measurements**

**TMS Apparatus and Stimulation Parameters**

Biphasic TMS pulses were applied using the MagVenture X100 stimulator (MagVenture A/S, Farum, Denmark) and a figure-of-eight coil (MC-B70, inner radius 10 mm, outer radius 50 mm). The maximum output of this coil and stimulator combination is approximately 1.9 T and 150 A/μs. The coil was manually held tangentially to the skull with the coil handle oriented perpendicular to the to be stimulated gyrus using the online visualization function of the BrainVoyager TMS Neuronavigator. The average Euclidean distance from our TMS coil to our 3 target sites was 2.7 cm. The estimated spatial resolution of the here used MagVenture MC-B70 coil at this distance is several cm³ (see Thielischer and Kammer 2004). Chronometric triple-pulse TMS was applied with an interpulse interval of 25 ms (40 Hz) at 120% resting motor threshold (MT). This event-related TMS protocol was expected to interrupt (strongly inhibit, not facilitate) the neuronal aspect for overt naming associated with the specific cortical region of interest that was targeted.

**TMS Localization**

Broca’s area is typically defined in terms of the pars opercularis and pars triangularis of the IFG, corresponding to areas 44 and 45 in Brodmann’s cytoarchitectonic map (Brodmann 1909). We targeted a site that was located superior to the apex of the vertical ascending ramus, which is thought to be the classical anatomical division for separating pars opercularis from pars triangularis.

Wernicke’s area is usually defined as the posterior 2/3 of the STG, posterior to Heschl’s gyrus (Naeser et al. 1987; Naeser and Pulombo 1994). For stimulation, we aimed for the posterior part of the left STG. This can also be described as the posterior part of Brodmann area (BA) 22 (Brodmann 1909).

The final region, according to Indefrey and Levelt (2004) responsible for conceptual preparation and lexical selection, was the midsection of the left MTG. This stimulation site, as well as IFG and posterior STG, was localized using the BrainVoyager TMS Neuronavigator. We thus used MRI-guided TMS to several anatomically defined network modules within the left hemisphere. This enabled us to account for individual differences in anatomical brain structures when stimulating. This approach was favored over a functional defined approach (e.g., **functional magnetic resonance imaging (fMRI)-guided TMS** since the specific psycholinguistic model we aimed to empirically test in our current study suggests a detailed description of both the anatomical location and time course of cerebral activations during speech production (Indefrey and Levelt 2004). This model exactly describes the respective critical anatomical sites during speech production that we used as target sites for our MRI-guided TMS. For precise localization of target points on each individual participant, see Figure 3. The stimulation order of these sites was randomized across participants.

**TMS Procedure**

Individual MTs were determined as the intensity at which the stimulation of the left motor cortex with single-pulse TMS resulted in a visible movement of the resting contralateral thumb in 50% of the trials. The MTs of the participants ranged from 27% to 42% of maximum stimulator output (mean = 35.40% [51.6 A/μs], SD = 4.6). The mean stimulation intensity was set at 120% of the MT and therefore resulted in 42.5% (65.75 A/μs) of maximum stimulator output (range 33-50%, SD = 5.3). Throughout the entire experiment, participants were wearing earplugs to protect their ears from the clicking sound and to minimize the interference of the sounds during the task.

Participants were tested with chronometric triple-pulse TMS in 3 separate sessions. Prior to starting with the experiment, participants were familiarized with the stimuli and practiced naming the stimuli repeatedly to reach a stable performance level in naming latency. Each experimental session consisted of 60 trials, divided into 4 blocks of 15 trials each. Stimuli were presented, and pulses were triggered using the software package “Presentation” (http://nbs.neuro-bs.com).

Chronometric TMS was applied at 5 different points in time following picture presentation onset, namely at 1) 150–175–200 ms, 2) 225–250–275 ms, 3) 300–325–350 ms, 4) 400–425–450 ms, and 5) 525–550–575 ms. In a sixth condition, no TMS pulses were applied during the trial. The presentation of the pictures, the TMS time window conditions, and the order of stimulation site were fully randomized across trials within each session.

**Figure 3. Localization of target points for each individual participant.** Surface reconstruction based on the WM-GM boundary of each individual participant with target points of stimulation. IFG displayed in green, MTG displayed in yellow, STG displayed in red. Standardized mean coordinates in Talairach space for each stimulation site were the following: IFG: \(x = -49\), \(y = 13\), \(z = 26\); MTG: \(x = -57\), \(y = -44\), \(z = 18\); STG: \(x = -55\), \(y = -44\), \(z = 18\).
The no TMS condition trials were randomly intermixed and thus included in our active TMS trials at the 5 different TMS time points, thereby controlling for many of the environmental nonspecific influences on task performance. Yet, the no-TMS condition does not control for known TMS-dependent nonspecific effects, such as the auditory and somatosensory stimulation, pain, muscle twitching, enhanced expectation, etc. We nonetheless decided against a fourth nonexperimental stimulation site (such as vertex or any cortical region that should not be involved in language production) as an additional control, since the specificity of our TMS effects are inherently controlled in the chronometric nature of the expected effects per site. In other words, since according to our hypothesis, we expect a time-specific TMS effect per site, including a significant interaction between target site \( \times \) time window, the above-mentioned TMS-dependent nonspecific side effects cannot account for any revealed time-specific differences in TMS affectivity across sites. In this sense, our 3 experimental target sites and their expected difference in temporal contribution represent a more appropriate control for our spatiotemporal hypotheses as a nonexperimental control site, such as vertex.

Chronometric TMS is here defined as an event-related TMS design where short bursts of TMS are applied online and time-locked to the stimulus event, providing us with an idea of the function of the region in a “time range” rather than a specific “time point.” This represents a compromise between the needed temporal resolution (hundreds of ms), the expected effect size of TMS, and the interference strength in case of higher cognitive functions with underlying highly distributed brain networks. It is by now common practice to apply, for example, short high-frequency TMS triplet bursts in chronicometric TMS studies of higher cognitive functions (see, e.g., Sack et al. 2005). It has been repeatedly and consistently shown that time locking such high-frequency bursts to the stimulus event provides reliable chronometric data regarding the relative critical time periods at which a certain brain region is functionally critical for a given task. However, this chronometric design increasingly runs the risk of possibly not only interacting with one network module but rather activate/inhibit different modules by means of such short rTMS trains. We ensured that there were no carryover effects between trials by testing for possible order effects and by carrying out respective pilot measurements with different intervals between trials (data not shown).

Data Analysis

Two of the participants did not go through the entire experimental TMS session since they experienced discomfort due to strong contractions of face muscles and were therefore excluded from the analysis. Incorrect trials of the remaining 10 participants were also excluded from the analysis. Incorrect trials (errors) were defined as semantically incorrect responses, hesitations, and extremely delayed responses (responses taking longer than 2000 ms). These incorrect trials constituted 3.77% of the original data acquired across the 10 participants.

The effect of TMS on the 3 different areas was tested at the above-mentioned 5 time intervals between stimulus onset and the TMS pulse, ranging from 150 to 525 ms following picture presentation. The RT data of the correct responses were further tested for normal distribution and variance homogeneity. These tests revealed that the RT data were positively skewed. In order to obtain a normal distribution, the entire data set underwent a logarithmic LN transformation. This ensured the suitability of the RT data for parametric statistical testing. Moreover, response times that were above or below 2 SDs of the mean were defined as outliers and were excluded from the analysis. These outlier trials constituted 3.06% of the original data acquired across the 10 participants.

Results

Error Analysis

Delayed responses and errors were analyzed together. We performed a 2-factorial repeated measures analysis of variance (ANOVA) with stimulation site and time window as within-subject factors on the error rate. Neither the main effect of site \((P_{2,18} = 0.85; P=0.44)\) nor the main effect of time window \((P_{4,45} = 0.15; P=0.98)\) or the interaction of the 2 factors \((P_{10,90} = 0.77; P=0.65)\) were significant, indicating that the amount of errors did not differ between stimulation sites and time windows.

TMS-Induced Changes in Picture Naming Latency

TMS over IFG (Broca’s area) showed time-specific effects on the reaction times during picture naming. Figure 4A shows that the RTs only slightly increased by 15 ms for the time window at 150 ms \((470 \text{ ms}; SD = 46)\) and by 16 ms for the time window at 225 \((471 \text{ ms}; SD = 43)\), as compared with the no TMS condition \((455 \text{ ms}; SD = 44)\). In contrast, the time window of 300 ms was characterized by a large effect of TMS on reaction times of approximately 50 ms \((504 \text{ ms}; SD = 75)\). At 400 and 525 ms, however, RTs rapidly decreased again and went back to baseline (no TMS) level. This indicates that TMS over Broca’s area had a very time-specific effect on picture naming, only interfering with behavior when applying triple pulse (tp) TMS 300 ms after picture presentation onset.

Stimulation of the midsection of the MTG led to different results. Reaction times in the no TMS condition were comparable to those when stimulating Broca’s area \((461 \text{ ms}; SD = 41)\) (see Fig. 4B). Applying TMS at 150 ms after picture onset led to a slight increase in reaction time \((484 \text{ ms}; SD = 40)\). This increase became very apparent when applying TMS at 225 ms \((496 \text{ ms}; SD = 42)\). Stimulation at 300 ms after picture presentation, in turn, led to a drop in reaction time \((482 \text{ ms}; SD = 39)\), followed by a peak in reaction time when applying TMS at 400 ms \((505 \text{ ms}; SD = 40)\). In the time window of 525 ms, RTs were decreasing again, reaching a level comparable to the no TMS condition \((474 \text{ ms}; SD = 39)\). These results suggest that applying tpTMS over MTG did affect picture naming at 2 points in time during the naming process, namely at 225 ms and 400 ms.

When stimulating Wernicke’s area, the no TMS condition was once again comparable to both the session in which Broca’s area was stimulated and the session in which the MTG was stimulated \((470 \text{ ms}; SD = 29)\). Interestingly, these RTs hardly varied when applying TMS at 150 ms \((471 \text{ ms}; SD = 44)\); at 225 ms \((473 \text{ ms}; SD = 33)\), at 300 ms \((474 \text{ ms}; SD = 49)\), and at 525 ms \((472 \text{ ms}; SD = 44)\) (see Fig. 4C). However, when applying tpTMS at 400 ms after picture presentation, an increase in reaction time of approximately 25 ms became visible \((495 \text{ ms}; SD = 22)\). These results indicate that TMS over Wernicke’s area had a very time-specific effect on picture naming, only interfering with behavior when applying tpTMS 400 ms after picture presentation onset.

Overall, in Figure 4, the effect of TMS on the 3 different stimulation sites can nicely be compared. While Broca’s area is the only stimulation site leading to an increase in RT at 300 ms, stimulation of Wernicke’s area led to an increase of RT only at 400 ms, whereas MTG stimulation led to a general increase in reaction time from 150 up to 400 ms, having a double peak at 225 and at 400 ms.

In order to test whether the time- and stimulation site-specific changes in reaction times also reached statistical significance, we analyzed the picture naming latency data based on the full factorial model according to our experimental design, thus performing a 2-factorial repeated measures ANOVA with stimulation site (IFG, MTG, STG) and time
TMS condition, applying TMS had a significant effect on simple contrasts analyses showed that, compared with the no stimulation condition in which no pulses were applied. This revealed a main effect of time window (F_{10,90} = 3.20; P < 0.01), showing that the time-specific effect of TMS differed between the stimulation sites. This significant interaction statistically supports the notion that the difference between the different time windows of TMS application is significantly different between stimulation sites, or in other words, that stimulation sites significantly differ in the time points at which neural activity is functionally relevant for successful picture naming.

This significant interaction term justifies additional site-specific analyses of IFG, STG, and MTG stimulation sessions.

When stimulating IFG, a main effect of time window was revealed (F_{4,18} = 0.86; P = 0.44) and a significant main effect of time window (F_{4,15} = 7.23; P < 0.0001), indicating that the effect of TMS differed between the various time points of stimulation. Importantly, the analysis also revealed a significant interaction of stimulation site and time window (F_{10,90} = 3.20; P < 0.01), showing that the time-specific effect of TMS differed between the stimulation sites. This significant interaction statistically supports the notion that the difference between the different time windows of TMS application is significantly different between stimulation sites, or in other words, that stimulation sites significantly differ in the time points at which neural activity is functionally relevant for successful picture naming.

This significant interaction term justifies additional site-specific analyses of IFG, STG, and MTG stimulation sessions.

When stimulating IFG, a main effect of time window was revealed (F_{4,15} = 5.55; P < 0.001). Simple contrasts were performed to compare the 5 conditions in which stimulation was applied at different time windows, with the baseline condition in which no pulses were applied. This revealed a significant difference in response time only for the condition in which stimulation took place at 300 ms after stimulus presentation (F_{1.9} = 9.04; P < 0.05; see Fig. 4A).

When stimulating the MTG, also a main effect of time window was revealed (F_{4,15} = 5.13, P < 0.01). Simple contrast analyses demonstrated a significant difference in reaction times between the no TMS condition and the time window of 225 ms (F_{1.9} = 1.77, P < 0.01) and between the no TMS condition and the time window of 400 ms (F_{1.9} = 29.45, P < 0.001; see Fig. 4B).

The one-factorial ANOVA of Wernicke’s stimulation also revealed a main effect of time window (F_{4,15} = 2.59; P < 0.05). Simple contrasts analyses showed that, compared with the no TMS condition, applying TMS had a significant effect on reaction times only in the time window of 400 ms (F_{1.9} = 13.6; P < 0.01; see Fig. 4C).

Discussion

The current study provides first direct empirical evidence that intact neural activity within the left IFG (Broca’s area), left posterior STG (Wernicke’s area), and the midsection of the left MTG, is functionally relevant and thus causally related to successful speech production. Hence, by using individualized MRI-guided chronometric TMS over all 3 regions in an within-subject design, we were able to show, for the first time under controlled experimental conditions, that left IFG, posterior STG, and the midsection of the left MTG all represent functionally relevant nodes of a widely distributed specific neurocognitive brain connectivity network underlying successful overt picture naming.

This study also showed that despite the question of functional relevance per se, online event-related TMS is also capable of charting the exact time point at which neural activity in a given brain region is critical for successful task performance. By applying such a TMS paradigm over several nodes of the same widely distributed brain network underlying speech production, we charted the relative time points of functional necessity in each of these network nodes, documenting a certain temporal order of functional relevance between distinct brain regions. This finding clearly indicates a specific spatiotemporal organization within the speech production network in terms of relative time course with each area contributing at different stages during the speech production process, suggesting distinct underlying functional roles within this complex multicomponential skill.

Concretely, we could show that left MTG is relevant at 2 distinct time points during picture naming, namely at an early stage and again at a later stage during speech production, as documented in our data by a second peak of functional relevance in left MTG. This second peak at around 400 ms temporally coincides with the functional relevance of posterior STG (Wernicke’s area). In contrast, IFG (Broca’s area) seems to be functionally relevant between the early and late MTG activity and thus slightly prior to the late functional relevance of posterior STG (Wernicke’s area), which occurred at the same time as the late MTG activity.

Thus, our study clearly revealed the functional relevance and causal relationship between intact neural activity in IFG, posterior STG, and the midsection of the left MTG for successful speech production, and our findings moreover clearly demonstrated that these 3 brain areas significantly contribute to successful speech production at different temporal stages. However, we want to point out that one needs to remain careful in interpreting the concrete temporal profiles with regard to concrete cognitive labels or models of speech production. Nonetheless, it has to be noted that with regard to the concrete predictions of the neurocognitive model of speech production presented by Indefrey and Levelt (2004), which we aimed to empirically test in the current study, our empirical data do not agree in all aspects with these predictions. According to the model, left MTG is functionally relevant first for lexical retrieval. This prediction is still in

Figure 4. Mean RTs (naming latencies) in ms (plus standard errors) for all time windows. Single asterisk indicates a significant difference of P < 0.05, double asterisk indicates a significant difference of P < 0.001. (A) TMS results of IFG interference. (B) TMS results of interference of the midsection of the left MTG. (C) TMS results of posterior STG interference.
accordance with our findings. In contrast to our findings, however, the model also predicts that Wernicke's area is relevant prior to Broca's area, underlying the cognitive subprocess of phonological code retrieval. Moreover, no second (late) functional relevance of MTG as revealed in our study is discussed in current speech production models.

In the following, we would like to make an attempt in interpreting and integrating our empirical findings with the existing literature and different models of speech production in order to account for this partial discrepancy. Moreover, we would like to offer possible alternative (maybe additional) functional roles of Broca's area, Wernicke's area, and MTG during speech production. It should be noted that it naturally becomes a matter of speculation and interpretation to post hoc assign a specific functional role to each of the stimulated brain regions based on our empirical findings (reverse-inference; [Poldrack 2006]) at this point, but we do believe that such a speculation is appropriate and useful.

According to the speech production model by Levelt et al. (1999), the first step in speech production planning is called "conceptualization." In this phase, the content of an utterance is represented as prelinguistic units or concepts. During the next step, called "formalization," concepts become lexicalized, that is, lexical entries corresponding to the concepts are retrieved from the mental lexicon. Formalization can be divided into 2 separate processes, namely "grammatical" (or syntactic) and "phonological encoding." During grammatical encoding, the syntactic structure of an utterance is specified. In contrast, during phonological encoding, the phonological form or sound of a word is specified (e.g., the phonemes or segments and the lexical stress) and so-called "phonological words" are created. After formalization is completed, each phonological word has to be converted into a format that can be used to control neuromuscular commands necessary for the execution of articulatory motor movements. The phonological word forms the basis for the retrieval of precompiled articulatory motor programs from a mental syllabary (Levelt and Wheeldon 1994; Cholin et al. 2006). These motor programs may be represented in terms of gestural scores, which specify the relevant articulatory gestures and their timing. The final step includes the execution of these gestures by the articulatory apparatus, which results in overtly produced speech.

We propose that the early effect at 225 ms in left MTG could represent the early process of lexical retrieval during which concepts become lexicalised (Salmelin et al. 1994). This is supported by various neuropsychological as well as noninvasive brain stimulation studies that suggested MTG and anterior lobe structures to play a role in conceptualization (Pobric et al. 2009; Schwartz et al. 2009; Lambon Ralph et al. 2010; Gallate et al. 2011). Regarding the functional relevance of IFG at around 300 ms, we argued on the basis of the results of an earlier study that the process being disturbed at this time point is likely to be the process of syllabification (Schuhmann et al. 2009). However, although other TMS studies over IFG have similarly and consistently revealed its functional importance for speech production and processing (Mottaghy et al. 1999, 2006; Shapiro et al. 2001; Sakai et al. 2002; Devlin et al. 2003; Nixon et al. 2004; Naeser, Martin, Nichols, Baker, Seekins, Kobayashi, et al. 2005; Andoh et al. 2006), our current data, and especially the relative timing of the functional relevance of IFG with regard to posterior STG and the left MTG, makes the exact functional contribution of IFG during the process of overt
of Broca's area has been observed to activate in semantic processing, whereas the left pars opercularis, relatively more in phonological processing. In a similar vein, pars opercularis likely also has a different WM trajectory to "posterior language zones," namely via arcuate fasciculus (AF) to anterior supramarginal gyrus, whereas the pars triangularis connects mainly via the extreme capsule to MTG and STG (for review, see Naeser et al. 2010).

Assuming that the early MTG effect at 225 ms represents the subsequent phonological encoding, the second peak of functional relevance found in left MTG and the peak in posterior STG at around 400 ms possibly indicate that the phonologically encoded concept may then be back projected from IFG to the left MTG and at the same time also forward projected to posterior STG. We propose that this feedback likely represents the neural connectivity mechanisms underlying internal speech monitoring (Leuninger et al. 2004; Christoffels et al. 2007; Christoffels et al. 2011; Schiller et al. 2009), and it may be part of the "motor theory of speech perception," as posited by Liberman, already decades ago (Liberman and Mattingly 1985).

Most speakers produce numerous words per second, seemingly without effort or conscious control of the speaking process. Nevertheless, we constantly monitor our own speech output on aspects, such as content, grammaticality, fluency, and volume. Without monitoring, producing speech can potentially lead to embarrassment, for instance, when taboo words are uttered unintentionally (Motley et al. 1982) or speech output can result in awkward mishearing (Garnes and Bond 1980). In word production, all critical subprocesses, such as conceptual preparation, lexical and syntactic encoding, phonological encoding, and articulation (see Levelt et al. 1999), are likely to be subject to such internal monitoring mechanisms. According to one of the most influential models of speech production (Levelt et al. 1999), self-monitoring is a centrally controlled process with limited capacity, which evaluates the quality of the speech by means of the speech comprehension process. The speech comprehension system, used for understanding speech of others, also subserves verbal external self-monitoring. In a similar vein, the abstract phonological code is presumably used for internal self-monitoring. Thus, it has been proposed that at the level of phonological encoding, potential speech production errors are controlled for via internal self-monitoring processes (see Levelt et al. 1999; Postma 2000. Hartsuiker and Kolk 2001) during which information is delivered to the speech comprehension system (posterior STG), where it is parsed and then transferred to the verbal monitor. The verbal monitor compares the parsed speech and the intentions of the speaker to the linguistic standards.

This interpretation is in line with the relative timing and feedback projections between IFG, left MTG, and posterior STG, as revealed in our study. We concretely identified an early effect in left MTG supposedly being the neural substrate underlying the process of lexical retrieval, after which neural information is sent forward to IFG for subsequent phonological encoding. IFG in turn back projects the phonologically encoded concept to left MTG while at the same time forward projecting it to posterior STG, thus the speech comprehension system, for internal self-monitoring purposes. Studies documenting the existence of direct and effective WM fiber connections between Broca's area and Wernicke's area date back to 1895 when Dejerine (1895) defined the AF as the prominent fiber tract connecting these 2 areas based on postmortem dissections. More recently, diffusion tensor imaging studies empirically identified and described these WM connections in the healthy living brain (Basser et al. 1994, 2000; Makris et al. 1997, 2005; Catani et al. 2002). These studies indicate that the AF is not the only WM tract connecting Broca's area and Wernicke's area but that there are additional dorsal and ventral pathways. Major connections from premotor cortices in the left hemisphere have been examined to follow a more "dorsal" route via the AF/superior longitudinal fasciculus III to the supramarginal gyrus (Croxonson et al. 2005; Frey et al. 2008; Saur et al. 2008); whereas major connections from ventrolateral prefrontal cortex (including pars triangularis) pursue a more "ventral" route via extreme capsule, to part of the STG or MTG (Frey et al. 2008; Saur et al. 2008). Separate dorsal and ventral pathways connecting parts of Broca's area with posterior language zones in the left hemisphere have also been suggested (Parker et al. 2005; Rushworth et al. 2006). The dorsal route in the left hemisphere, as summarized by Frey et al. (2008), is largely restricted to sensory-motor mapping of sound to articulation and higher order articulatory control of speech, where the pars opercularis is connected directly with premotor area 6 (involved with orofacial musculature) (Petrides et al. 2005). The vPMC was observed to have connections with the horizontal portion of the AF in both the left hemisphere and in the right hemisphere, similar to the pars opercularis in each hemisphere (Kaplan et al. 2010). Thus, both the pars opercularis and the vPMC are thought to be connected with the anterior supramarginal gyrus via the dorsal route in each hemisphere. The ventral route in the left hemisphere, however, likely performs linguistic processing of sound to meaning, requiring tempo-frontal interaction and top-down regulation of linguistic processing such as that involved in verbal retrieval and lexical/semantic aspects of language processing (Price et al. 1996; Poldrack et al. 1999; Gold and Buckner 2002; Devlin et al. 2003; Nixon et al. 2004; Saur et al. 2008). It has also been suggested that the pars triangularis is more related to verbal fluency in general and not restricted to semantic fluency (Heim et al. 2008). Gough et al. (2005) also provide support for the notion that there must be a dissociation between the roles of left pars triangularis versus left pars opercularis in semantic versus phonological tasks, by applying TMS to these 2 areas in normals, with differential/opposite effects observed (for review, see Naeser et al. 2010).

Our data suggest that the existence of such WM pathway connections between these prominent language-related brain regions, such as IFG, MTG, and posterior STG, might be of particular functional relevance during speech production and that the human language function is thus not only based on the GM of circumscribed brain regions in the frontal and the temporal cortex, but that instead, successful speech production largely depends on intact WM fiber tracts connecting these adjacent as well as distant language-related brain regions. However, it remains to be resolved which exact parts of Broca's...
area (pars opercularis vs. pars triangularis) and which WM pathways (AF vs. extreme capsule) to posterior language zones (anterior supramarginal gyrus area vs. middle and/or STG areas) are involved during speech production and when they are involved. Based on our current findings, follow-up eTMS studies could be designed to more specifically target 1) the pars opercularis portion that is closer to the vPMC (likely relevant for phonological encoding and syllabification, after conceptualization from the MTG), 2) the anterior supramarginal gyrus (likely related to the pars opercularis/vPMC in forming the timing for a dorsal route), and 3) the pars triangularis portion that is located further away from the pars opercularis (with direct WM connection to MTG and STG, via the extreme capsule, for a feedback loop for top-down semantic processing, forming the timing for a ventral route).

The given interpretation and cognitive labeling of our findings assumes that the revealed relative timing differences in functional contribution are indicative of a temporal sequence of information flow in which one area processes a certain aspect of the task and subsequently sends this information to another brain region for further processing. It is important to note that such interpretation of effective brain connectivity based on chronometric TMS data is by definition implicit. The chronometric TMS results do not provide direct evidence of feed forward or feedback backward flow of information, but they do show timing differences in functional relevance between brain regions within one network, thereby providing strong but indirect evidence regarding neural information flow.

To empirically complement such evidence for a particular effective connectivity model, it would be most interesting and useful to cross-evaluate with some effective connectivity models by combining TMS chronometry with the evidence using, for example, structural equation modeling or dynamic causal modeling and/or Granger causality mapping (see, e.g. de Graaf et al. 2009).

Funding

Notes
We thank our medical supervisor Cees van Leeuwen and our independent physician Martin van Boxtel. We thank Joel Reithler for the support in figure preparation. We also thank the anonymous reviewers for their very constructive comments. Conflict of Interest: None declared.

References