Behavioral/Cognitive

Low-Frequency Oscillations Code Speech during Verbal Working Memory

Johannes Gehrig,[1] **Georgios Michalareas,**[2] **Marie-Therese Forster,**[3] **Juan Lei,**[1,4] **Pavel Hok,**[1,5] **Helmut Laufs,**[1,6] **Christian Senft,**[3] **Volker Seifert,**[3] **Jan-Mathijs Schoffelen,**[7] **Simon Hanslmayr,**[8] **and Christian A. Kell**[1]

[1]Department of Neurology, Goethe University, 60528 Frankfurt, Germany, [2]Max Planck Institute for Empirical Aesthetics, 60322 Frankfurt, Germany, [3]Department of Neurosurgery, Goethe University, 60528 Frankfurt, Germany, [4]Institute for Cell Biology and Neuroscience, Goethe University, 60438 Frankfurt, Germany, [5]Department of Neurology, Palacky University and University Hospital Olomouc, 77147 Olomouc, Czech Republic, [6]Department of Neurology, Christian-Albrechts-University, 24105 Kiel, Germany, [7]Radboud University, Donders Institute for Brain, Cognition and Behaviour, 6525 HR Nijmegen, The Netherlands, and [8]School of Psychology at University of Birmingham, B15 2TT Birmingham, United Kingdom

The way the human brain represents speech in memory is still unknown. An obvious characteristic of speech is its evolvement over time. During speech processing, neural oscillations are modulated by the temporal properties of the acoustic speech signal, but also acquired knowledge on the temporal structure of language influences speech perception-related brain activity. This suggests that speech could be represented in the temporal domain, a form of representation that the brain also uses to encode autobiographic memories. Empirical evidence for such a memory code is lacking. We investigated the nature of speech memory representations using direct cortical recordings in the left perisylvian cortex during delayed sentence reproduction in female and male patients undergoing awake tumor surgery. Our results reveal that the brain endogenously represents speech in the temporal domain. Temporal pattern similarity analyses revealed that the phase of frontotemporal low-frequency oscillations, primarily in the beta range, represents sentence identity in working memory. The positive relationship between beta power during working memory and task performance suggests that working memory representations benefit from increased phase separation.

Key words: electrocorticography; memory representations; sentence repetition; speech perception; speech production; temporal pattern similarity

---

Significance Statement

Memory is an endogenous source of information based on experience. While neural oscillations encode autobiographic memories in the temporal domain, little is known on their contribution to memory representations of human speech. Our electrocortical recordings in participants who maintain sentences in memory identify the phase of left frontotemporal beta oscillations as the most prominent information carrier of sentence identity. These observations provide evidence for a theoretical model on speech memory representations and explain why interfering with beta oscillations in the left inferior frontal cortex diminishes verbal working memory capacity. The lack of sentence identity coding at the syllabic rate suggests that sentences are represented in memory in a more abstract form compared with speech coding during speech perception and production.

Introduction

Rhythmic regularities in the physical environment entrain neural oscillations in the brain. Such a temporal relationship is well described for human speech perception but could result from mere rhythmic stimulus-induced evoked activity. Here we make use of the unique human ability to represent speech in working memory, to study if and how neural oscillations code speech endogenously during verbal memory processes.

Covertly remembering and overtly repeating speech require short-term verbal working memory (WM) (Baddeley, 2003; Jacquemot and Scott, 2006; Bastiaansen et al., 2010; Perrone-

Table 1. Patient characteristics

Patient no.	Gender	Age(yr)	Tumor localization	Diagnosis	Native language
1	Male	23	Frontal	Diffuse astrocytoma, WHO II	German
2	Male	53	Postcentral	Anaplastic astrocytoma, WHO III	German
3	Male	33	Frontal	Anaplastic astrocytoma, WHO III	English and Russian
4	Male	63	Temporal	Glioblastoma, WHO IV	German
5	Male	34	Frontal	Glioblastoma, WHO IV	English and Dutch
6	Male	54	Frontal	Anaplastic astrocytoma, WHO III	German
7	Female	34	Temporal	Anaplastic glioma, WHO III	German
8	Male	27	Temporal	Anaplastic glioma, WHO III	German
9	Female	27	Fronto-opercular	Pilocytic astrocytoma, WHO I	German

Bertolotti et al., 2014; Alderson-Day and Fernyhough, 2015) as well as a transformation of auditory into motor speech representations (Hickok and Poeppel, 2000, 2004; Cogan et al., 2014, 2017; Cheung et al., 2016). One classical view in theoretical WM models is such that WM relies in part on a left-lateralized phonological loop involving interactions between frontal and temporal cortical areas (Baddeley, 2003; Jacquemot and Scott, 2006; Hickok and Poeppel, 2007; Buchsbaum and D'Esposito, 2008; Herman et al., 2013), where perceived speech is represented in left temporal cortex, and articulatory motor programs are represented in left frontal cortex. Recent observations during syllable repetition, using left perisylvian electrocorticography (ECoG), document distinct brain regions that store phonological input, transform sensory into motor representations, and maintain the motor output in memory (Cogan et al., 2017). Yet, in natural speech, syllables build up words that are used to construct sentences. Sentences are more easily remembered compared with (pseudo)word lists because of their syntactic structure and sentence-level semantics, which logically connect single words. This suggests additional processes that could facilitate sentence compared with word list recall. Association cortices in the frontal and temporal lobe contribute to this effect (Bonhage et al., 2017) and may mediate interactions between WM systems and higher-order sentence-level representations (Potter and Lombardi, 1990, 1998). Yet, how the brain codes natural speech in WM is unknown.

Neural oscillations are a good candidate for representing speech in WM because they play an important role in memory formation (Buzsáki, 2006); and during speech perception, neural oscillations are entrained by the exogenous quasi-rhythmic nature of the auditory speech signal (Giraud and Poeppel, 2012). The alignment of the oscillatory neural activity with the temporal structure of the sensory input at overlapping frequencies facilitates the temporal parsing of the speech signal and supports speech comprehension in a noisy environment (Luo and Poeppel, 2007; Nourski et al., 2009; Mesgarani and Chang, 2012; Peña and Melloni, 2012; Golumbic et al., 2013; Gross et al., 2013; Kubanek et al., 2013; Ding and Simon, 2014; Rimmele et al., 2015). In the auditory association cortex, the phase of neural oscillations in the theta band (4–8 Hz), which approximately corresponds to the syllabic modulation rate of the speech signal, and interactions between theta phase and power in the high beta/low gamma (25–35 Hz) and high gamma band (60–80 Hz) have been associated with speech coding during speech perception (Giraud and Poeppel, 2012). A computational speech perception model proposes that the auditory cortex chunks beta-cycle long time windows of the quasi-continuous speech signal for template matching with speech representations in memory (Ghitza, 2011). Oscillations in the beta range have been proposed to aid sentence unification (Bastiaansen et al., 2010).

Information coding may be complex because, even during speech perception, neural oscillations in motor cortex can "represent" speech information on the basis of acoustic features rather than in a "motor type" organization on the basis of somatotopic maps (Cheung et al., 2016). This challenges classical views of structure–function relationships and information representation. Consequently, verbal WM models have to be revisited. Using ECoG data, we thus investigated whether and how neural oscillations code speech during sentence reproduction, specifically the identity of the repeated sentence during WM. Our results reveal that the phase of endogenous low-frequency oscillations, particularly in the beta band, code speech information, even in the absence of sensory input during verbal WM.

Materials and Methods

Participants. Perisylvian ECoG of the left language-dominant hemisphere was recorded during awake tumor surgery in 2 female and 7 male right-handed patients (Table 1). Evaluation of handedness was based on self-report and patient observation during the perioperative weeks. Our sample was restricted to patients with left perisylvian craniotomies because we did not operate on right language-dominant patients during the course of the study. All patients were left language-dominant, as nonmotor speech arrests and anomia sites were detected by direct cortical electrical stimulation of left perisylvian cortex. In a clinical setting, ECoG can be used to guide tumor resection and monitor epileptic discharges. The ECoG session that we report here was research driven. All participants gave written and informed consent, and the study was approved by the local ethics committee (GZ 310/11).

Experimental design. After completion of the clinical language testing during direct cortical stimulation, participants performed a sentence reproduction task during ECoG recording. Participants listened to prerecorded three-word sentences (listening phase) of their own voice (normalized between samples to −3 db relative to full scale and presented via loudspeakers at a comfortable loudness level for the patient) with a maximum length of 1.5 s and waited for 1.5 s (maintenance phase) until visual presentation of a go cue for sentence reproduction (speaking phase; see Fig. 1A). This introduced a verbal WM maintenance phase into the study design.

Sentences were recorded using a Philips PC headset (SHM7410U) with an adjustable noise-canceling boom microphone (50–15,000 Hz, −42 dB, 2.2k Ohm) and recorded using Adobe Audition (RRID: SCR_015796) at a sampling rate of 44,100 Hz. Sentences were presented via a RAIKKO NANO Vacuum Speaker (2.5 W). The average distance to the patient's ear was 70 cm.

In 5 participants, a specific sentence (same sentence) was used in 32 trials, whereas 73 further trials were based on 73 different sentences. Using same and different sentences is a prerequisite for the temporal pattern similarity analyses described below. In 4 participants, we could increase the trial number to 60 times the same sentence and 65 times different sentences. The sentences had an identical syntactic and syllabic structure consisting of pronoun (1 syllable), verb (1 syllable), and adverb (2 syllables) to avoid syntactic differences affect the results (e.g., "Er rennt immer", "Du bist wichtig"). Consequently, the acoustic speech envelopes

of same and different sentences correlated with each other (median correlation coefficient of 0.26; *p* = 0.0039, Wilcoxon signed rank test). The 7 native German speakers were tested with a German corpus and 2 non-German early bilingual English speakers were tested using English sentences of the same structure.

Recording and preprocessing. ECoG data were acquired with high-resolution grids (5 mm spacing, Ad-Tech Medical), referenced against a frontocentral subgaleal needle electrode (Fz) and sampled at 5 kHz (BrainAmp MR plus amplifier, BrainProducts, RRID:SCR_009443). Grid dimensions (between 64 and 128 equally spaced electrodes) were limited by the size of the individual craniotomies. Synchronized video, EMG of orbicularis oris and orbicularis oculi muscles, and the patient's voice were recorded to identify speech onsets. For ideal synchronization between the ECoG recordings and the auditory data, the microphone output was also fed into the EMG amplifier. The FieldTrip toolbox (http://FieldTrip.fcdonders.nl/, RRID: SCR_004849) (Oostenveld et al., 2011) and in-house MATLAB code (MATLAB with Statistics and Signal Processing Toolbox Release 2012b, The MathWorks, RRID:SCR_001622) were used to preprocess and analyze the data. ECoG data were high-pass filtered at 1 Hz and low-pass filtered at 300 Hz using a hamming, two-pass Butterworth filter with a filter order of 5. The line noise was removed using band-stop filters for the 50 Hz noise and the harmonics up to 200 Hz using hamming, two-pass Butterworth filter with a filter order of 4. Data were rereferenced using bipolar montages to increase spatial resolution and to minimize the influence of global sources (Mercier et al., 2017). Vertical and horizontal montages of neighboring electrodes further increased spatial resolution. For a first analysis, the trial onsets were defined as the auditory stimulus onset. This analysis identifies effects time-locked to auditory stimulus onset. In a second analysis, the trial onset was defined as individually marked auditory stimulus offsets to focus on WM processes that could only start once the entire sentence was perceived. In a third analysis, effects time-locked to speech production were investigated by using individual speech onsets (voice onset) to define *t* = 0. The delay between the visual Go cue and the speech onset was defined as reaction time for analyses of the behavioral data. Wrong repetitions or lapses were labeled as error trials.

Electrodes over the radiologically defined tumor were excluded from analyses. Consequently, number of electrodes and covered brain regions varied between participants (Table 2). Trials containing artifacts (excessive noise, jumps, DC shifts) were manually removed from each dataset. Artifact trials were excluded, leaving 69.6% (±12.1%) trials for analyses. Of these, 34.4% (±15.2%) were same sentence trials and 65.6% (±15.2%) different sentence trials.

Electrode selection. Bipolar montages between directly adjacent contacts (hereafter electrodes) were selected based on their anatomical location and activation profile in the time-frequency spectra. Electrodes of interest were localized in mid and posterior superior temporal gyrus (STG), pars triangularis and pars opercularis of the inferior frontal gyrus (IFG), premotor cortex, and the dorsal and ventral primary motor cortex (see Fig. 1B). The activation profile for electrode selection was defined as a function of task-specific activity, time-locked to the onset of the auditory stimulus (regional cluster of suppression of low-frequency oscillations and positive gamma band responses; see Fig. 2). Time-frequency spectra were created using Slepian multitapers (2–170 Hz) implemented in the FieldTrip toolbox with a frequency resolution of 1 Hz and a frequency-adaptive spectral smoothing of 0.4 per frequency. The elec-
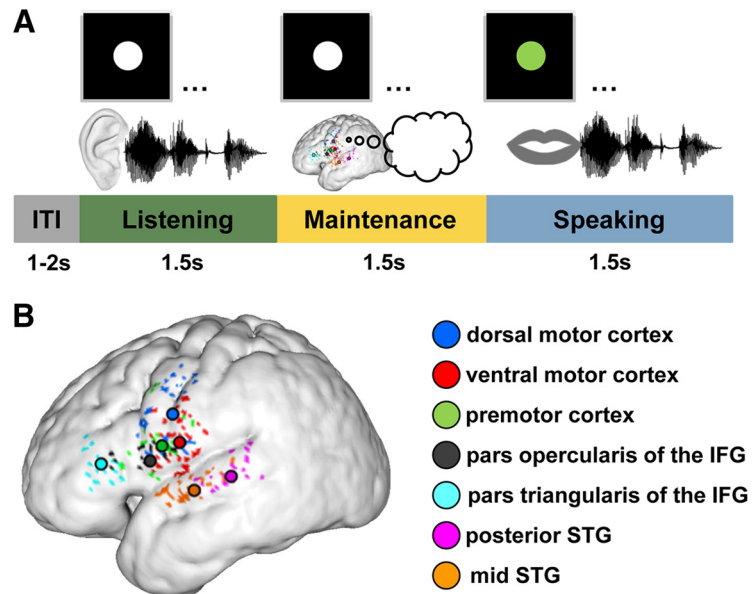

Figure 1. *A*, Sentence reproduction paradigm. Participants listened to three-word sentences while a white circle was presented on a screen. This listening phase (in all figures color-coded in green) was followed by a 1.5 s delay in which the sentence had to be maintained in WM (color-coded in yellow). Thereafter, the circle turned green for 1.5 s, which served as a go cue for sentence reproduction (speaking phase, color-coded in blue). The intertrial interval (ITI), which started once the circle turned white, was jittered in duration between 1 and 2 s (0.5 s steps). *B*, Selected electrodes on a group average reconstructed cortical surface. Individual ECoG sites were transformed nonlinearly into a common space and overlaid on top of a 3D reconstructed average cortical (i.e., pial) surface. Black-rimmed circles represent anatomical center of gravity of electrodes (colored dots) in dorsal motor cortex (blue), ventral motor cortex (red), premotor cortex (green), pars opercularis of the IFG (black), pars triangularis of the IFG (cyan), posterior STG (pink), and mid STG (orange). On the group level, electrodes intertwined; but in individual participants, dorsal motor cortex electrodes were always more dorsal than ventral motor cortex electrodes. In a minority of participants, the functionally defined clusters were located slightly more caudal than the group mean. As a consequence, some participants showed motor-related activity in the postcentral gyrus.

Table 2. Number of electrodes per region and participant

Participant no.	Ventral motor	Dorsal motor	Premotor	IFG pars triangularis	IFG pars opercularis	Mid STG	Posterior STG
1	2	—	2	—	2	4	5
2	5	5	2	—	—	4	—
3	7	4	5	6	6	5	5
4	6	—	3	—	3	6	6
5	6	7	2	5	3	—	—
6	4	—	2	4	—	—	—
7	5	6	—	—	3	3	2
8	3	3	3	2	2	2	—
9	3	5	3	—	2	3	—

trode with the largest power changes compared with baseline (−800 ms to −100 ms before stimulus onset), and its direct neighbors were defined as an ROI. The relationship between individual electrode clusters and anatomy served to identify regions in individual participants. The mean power over electrodes in a region was calculated and averaged over participants. As this specific time-frequency analysis only served to select electrodes of interest, no further statistics were performed on these time-frequency results. For visualization in Figure 2, the ColorBrewer RdBu colormap was used to avoid the induction of artificial perceptual boundaries (www.ColorBrewer.org; Cynthia A. Brewer, Department of Geography, Pennsylvania State University).

Electrode localization. All participants underwent a preoperative structural MRI session (Trio 3T scanner, Siemens) with a standard head coil. The MRI protocol included a high-resolution T1-weighted MPRAGE sequence (TR 2250 ms, TE 3.83 ms, partial Fourier 7/8, FOV = 256 × 224 mm, 144 slices, and isotropic voxel size 1.0 mm) for anatomical reference.
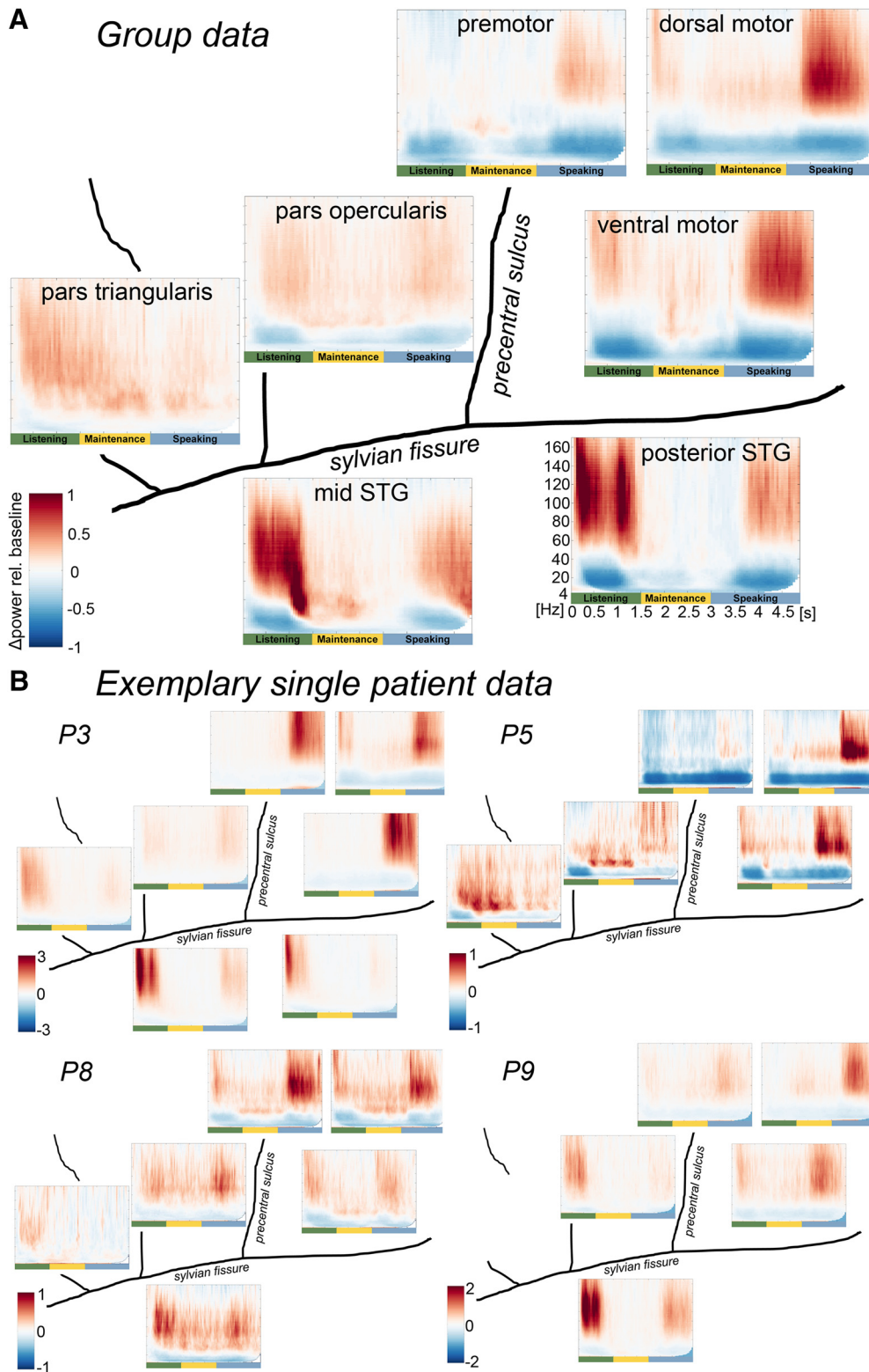
Figure 2. Group-averaged (**A**) and exemplary single patient (**B**) time frequency-resolved data of left perisylvian ECoG during sentence reproduction. All panels: Left to right, Rostrocaudal axis. Top to bottom, Dorsoventral axis. Plots represent averages over clusters of electrodes that share the same spectral pattern over time within an anatomical region. Power is color-coded as the relative change per cent compared with baseline. A value of 1 indicates a change of 100%. Frequency on the *y* axis (4 –160 Hz), time on the *x* axis (0 –5 s with trials locked to auditory sentence onset). Trial phases are illustrated in green (listening), yellow (maintenance), and blue (speaking).

The individual pial cortical surfaces, as well as cortical and brain masks of the left hemisphere were reconstructed from the T1-weighted structural image using the FreeSurfer (RRID:SCR_001847) standard pipeline implemented in the "recon-all" tool (Dale et al.,

1999). The mass lesions were manually masked during the preprocessing to correct local reconstruction errors. Electrodes were localized manually on each individual's pial surface using intraoperative photographs. The individual structural images were transformed

nonlinearly to a common template, which was linearly aligned to the MNI standard space (Grabner et al., 2006) using Dartel, part of SPM12 (SPM, RRID:SCR_007037) (Ashburner, 2007).

Electrode positions are illustrated in Figure 1B on the group average-normalized cortical surface. Rimmed circles represent the center of gravity of electrodes in a given cortical region.

Temporal pattern similarity (TPSim). To investigate temporal information coding of sentence identity during verbal WM, we computed TPSim (Staudigl et al., 2015; Michelmann et al., 2016) of the ECoG signal during sentence reproduction. TPSim is a form of representational similarity analysis (Kriegeskorte et al., 2008; Yaffe et al., 2014), a well-established methodology used particularly for investigating memory representations. In short, it is based on a correlation analysis of time courses of brain activity. In contrast to the more widely used intertrial coherence (e.g., Luo and Poeppel, 2007), it does not reflect the consistency of the signal's phase or amplitude across trials but rather detects the consistency of temporal modulations in the signal (Golumbic et al., 2013). TPSim thus represents the time-resolved correlation between the spectral coefficients of trial pairs within a given time window and a given frequency band. Our TPSim analyses reveal correlations between neural signals that reflect consistent processing in time.

Recent studies on speech perception have identified the phase of low-frequency oscillations as well as broadband high gamma power envelope fluctuations as temporal information carriers during stimulus-related processing (Luo and Poeppel, 2007; Mesgarani and Chang, 2012; Golumbic et al., 2013) and sentence processing (Nelson et al., 2017; Tang et al., 2017). Based on these studies, we hypothesized that both parameters could be involved in temporal information coding during verbal WM. We thus investigated both the complex coefficient of low-frequency oscillations' phase-locking value (phase TPSim) and the broadband high gamma power envelope fluctuations (power TPSim). For investigating low-frequency oscillations as carriers of sentence identity processing, the frequency range (4–48 Hz) was selected. In this frequency range, resolved at 1 Hz, time-frequency analysis of each trial was performed using wavelet analyses (Fieldtrip toolbox) in time steps of 10 ms and the resulting complex Fourier coefficients were used for the subsequent similarity analysis. To also investigate high gamma power envelope fluctuations as carriers of sentence identity processing, the trial data were band-passed in the frequency range (70–170 Hz) (Mesgarani and Chang, 2012; Golumbic et al., 2013; Mesgarani et al., 2014; Herff et al., 2015) using two-way least-squares FIR filtering and the absolute values of the Hilbert transform were computed.

For low-frequency oscillations (encoding sentence identity), time-resolved phase TPSim was computed in the following way. A sliding time window of 500 ms length in steps of 100 ms was used along the time bins for which spectral coefficients were computed (see above). In each of these time bins and for each frequency in the low-frequency range (4–48 Hz), all the spectral coefficients falling within the centered sliding time window were selected for each trial in the two conditions (same and different sentence trials). As we expected, the encoding to occur in the phase variations, rather than in the power of low-frequency oscillations (Luo and Poeppel, 2007; Mesgarani and Chang, 2012; Golumbic et al., 2013), the phase TPSim analysis was performed on phase variation similarities between trials rather than power variation similarities. Within each of the two conditions, for each possible pair of trials, a complex coefficient of phase-locking between the two spectral coefficient series was computed and a Fisher z transformation was applied. This coefficient is almost identical to the phase-locking value (PLV) (Lachaux et al., 1999) with the main difference that, instead of taking the magnitude of the complex phase-locking coefficient, to compute PLV, the complex coefficient itself is used. This retains the phase difference information for every pair of trials so that it is taken into consideration in the phase TPSim analyses. PLV, as well as the complex coefficient used here, is a measure very similar to coherence but with normalized cross-spectra so that the effect of amplitude variations is masked and only phase variations are examined. These analyses were repeated for each condition, each given time-frequency bin, and each electrode across all possible pairs of trials (except autocorrelations, see Formula 1) as follows:

$phaseTPSim(c,t)$

$$= \sum_{j=1}^{Ntc} \sum_{k=1}^{Ntc} N_w^{-1} \sum_{\tau=t-L}^{t+L} \exp(i(\varphi(j,\tau,c) - \varphi(k,\tau,c)))$$

$$where\ j \neq k\ and\ j,\ k\ are\ trial\ indices \quad (1)$$

where c indicates condition (i.e., same or different sentence), t indicates the time point at the center of the moving time window in a trial, L indicates the extent of the time window from its center, φ indicates the phase of the complex spectral coefficient for a given trial and time point, Ntc indicates total number of trials for the given condition c, and N_w indicates number of time points within the moving window equal to $2L + 1$. The term inside the exponent is a complex number expressing the phase difference between two trials at a given time point.

For high gamma power envelope fluctuations, time-resolved power TPSim was computed in the following way. A sliding time window of 500 ms length in steps of 100 ms was used along the time range for which the high gamma power envelope was computed (see above). In each of these time bins, all the high gamma power envelope values falling within the centered sliding time window were selected for each trial in the two conditions (same and different trials). Then, within each of the two conditions, for each possible pair of trials, the Pearson's correlation coefficient was computed between the power envelope series of these two specific trials from the given condition, time-frequency bin, and electrode. For each condition, each given time bin, and each electrode, the analysis was repeated across all possible pairs of trials; and after applying a Fisher z transformation, the mean of the resulting distribution of correlation values was computed as the metric to describe the overall similarity between trials and was assigned as the TPSim value for the specific case (see Formula 2).

$powerTPSim(c,t)$

$$= \sum_{j=1}^{Ntc} \sum_{k=1}^{Ntc} \frac{E[(h(j,t_w,c) - \overline{h(j,t_w,c)})(h(k,t_w,c) - \overline{h(k,t_w,c)})]}{\sigma(j,t_w,c)\sigma(k,t_w,c)} \quad (2)$$

where h indicates the absolute of the Hilbert transform for given trial, time, and condition, $\overline{h}$ indicates the mean of the Hilbert coefficients, E indicates expected value, and $\sigma(j, t_w, c)$ indicates SD of Hilbert coefficients within the time window centered at time t.

The term inside the sums describes the Pearson correlation coefficient. The expected value in the nominator and the SD values in the denominator are computed for all the Hilbert coefficients inside the time window centered at time t. The values within this window are represented in Equation 2 by the variable t_w as follows:

$$t_w = \{t - L, t - L + 1..., t, ..., t + L - 1, t + L\}$$

The mean of the resulting distribution of phase-locking (low frequencies) or power correlation values (high frequency) was computed as the metric to describe the overall similarity between trials and was assigned as the TPSim value for the specific case.

To test for temporally structured brain activity that encodes sentence identity, we compared TPSim of neural activation in trials when the same sentence was used with TPSim in the trials with different sentences, separately for phase and power TPSim (see Eq. 3). A difference in correlation of neural signals between same and different sentence trials (ΔTPSim) identifies sentence-specific temporal information coding, which is the central parameter of interest in our study. ΔTPSim was investigated across all time(-frequency) bins and electrodes. This resulted in a ΔTPSim per patient, electrode, and time(-frequency) bin. ΔTPSim was expected to be high when neural signals contain more similar temporal features in the case of same sentence trials, compared with different ones as follows:

$phase/power\Delta TPSim(t) = phase/powerTPSim(same,t)$

$$- phase/powerTPSim(diff,t) \quad (3)$$

To detect temporal information coding that is directly related to perceptual processing during listening, the trials were time-locked to the onset of the
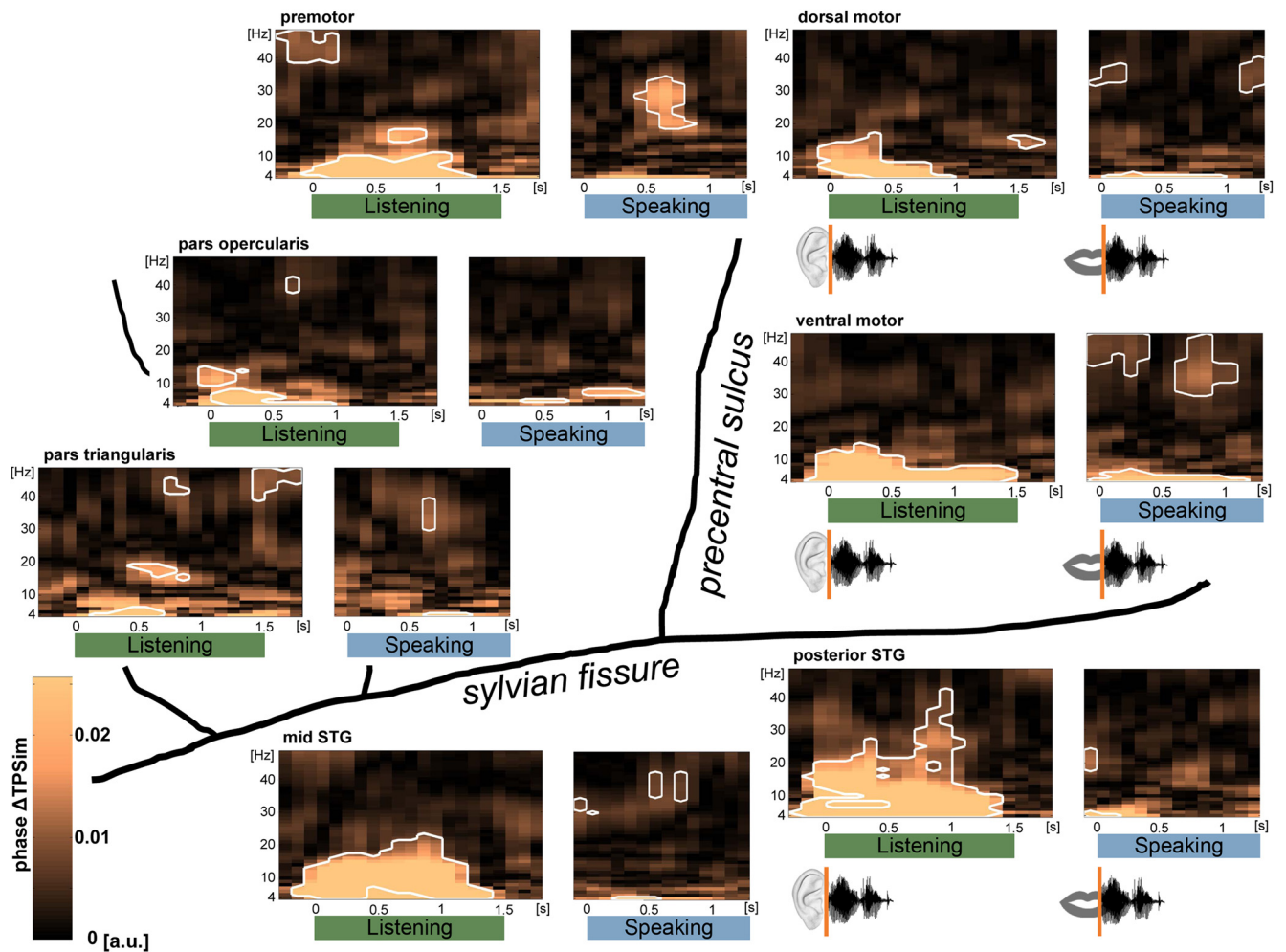
Figure 3. Perception- and production-related sentence identity coding in time (phase ΔTPSim): group data of left perisylvian ECoG. Plots represent averages over electrodes and participants. For perception-related temporal information coding during listening (green), trials were time-locked on stimulus onset; whereas for the analysis of production-related sentence identity coding during speaking (blue), trials were time-locked on individual speech onsets. Results of the analyses performed on the complex coefficient of phase-locking values of the time frequency-resolved data in the low-frequency range (4–48 Hz) are shown in each region. Significant clusters are framed in white ($p < 0.01$, stepwise Bonferroni-corrected). Sentence identity coding (phase ΔTPSim) was found in most frequencies almost instantaneously after stimulus onset in most of the studied frequencies and regions. During speaking, phase ΔTPSim was observed in relative higher frequencies and a narrow theta band. Time points on the x axes correspond to the center of the 500 ms analysis window.

auditory stimulus at the beginning of the listening phase. To investigate temporal information coding during speech production, the same analysis was performed with the data time-locked to individual speech onsets. Results from the phase ΔTPSim analyses are depicted in Figure 3.
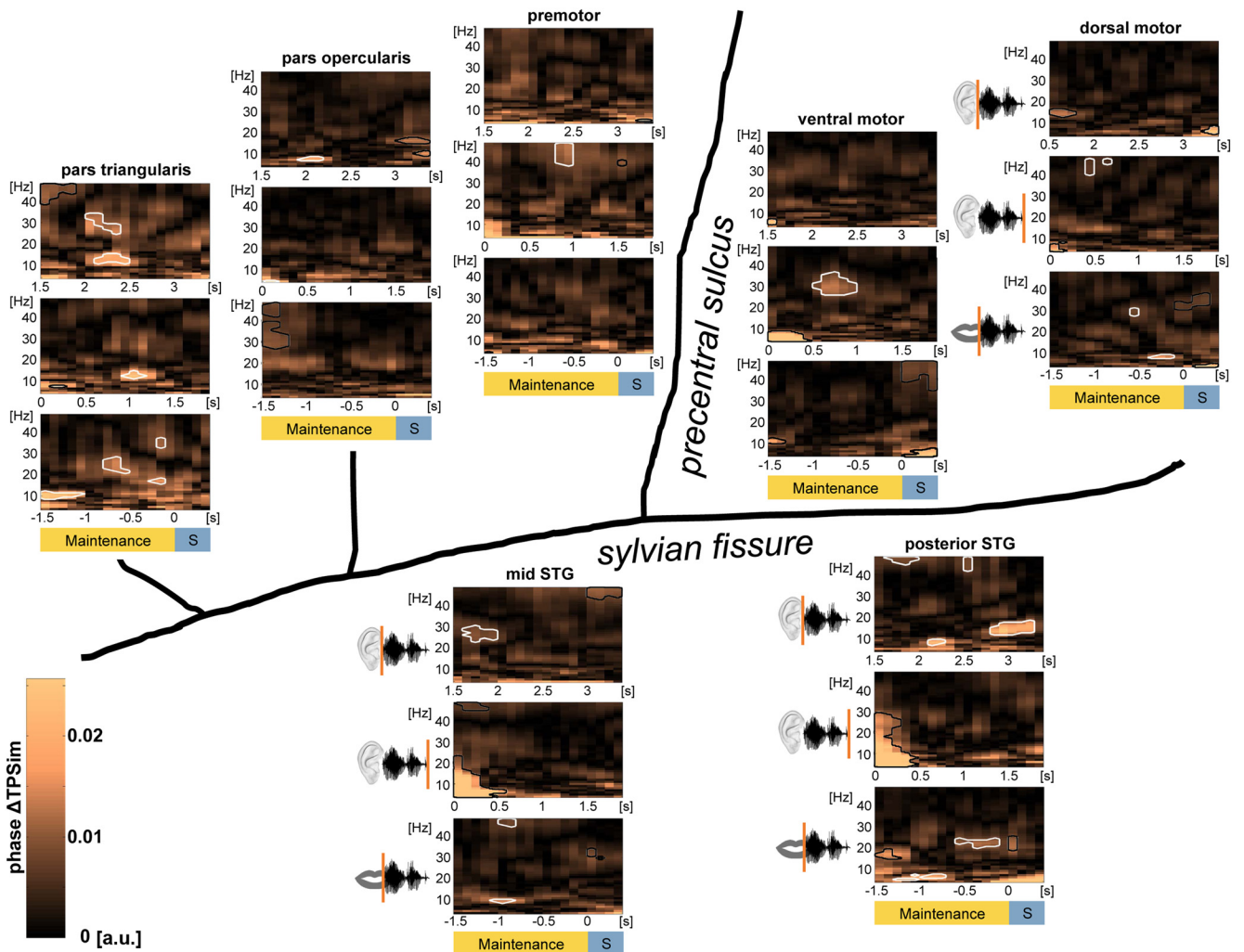
ΔTPSim during the WM maintenance phase was investigated in three separate analyses. Temporal information coding related with the recorded sentence onset reflecting sentence identity coding in WM was analyzed in the data cut on auditory stimulus onset (phase ΔTPSim; see Fig. 4, top). Processes that could start only after the entire sentence was perceived were detected in analyses in which the data were time-locked to the offset of the auditory stimulus (phase ΔTPSim; see Fig. 4, middle panels). Processes during WM maintenance that were more directly related with speech production than perception were investigated in analyses in which the data were cut on the individual speech onset times (phase ΔTPSim; see Fig. 4, bottom panels). Results of the phase and power TPSim analyses, separately, are represented in Figure 5.

We used ΔTPSim between same and different sentence trials to identify significant clusters of temporal information coding of sentence identity. The comparison of a repeated with nonrepeated items introduces repetition effects. Such priming effects were excluded from the sentence identity analyses by masking out repetition-related ΔTPSim. This was based on a comparison of TPSim of directly repeated versus nondirectly repeated same sentence trials (response priming) (Henson et al., 2014).

Additional adaptation effects (repetition priming) (Henson et al., 2014) were excluded based on a second comparison between same sentence trials in the second versus first half of the experiment (exclusive masking; see Statistical analysis). However, repetition priming is more closely related to memory representations on a larger time scale compared with response priming and is thus worth exploring during WM maintenance. Because Participant 1 did not contribute a sufficient number of repeated trials, this analysis was based on data from the remaining 8 participants. These additional analyses were performed on low-frequency phase and broadband gamma power cut on the auditory stimulus onset and offset and on the speech onset (phase ΔTPSim, see Fig. 6; power ΔTPSim reported in the text). Results of the power ΔTPSim analyses are illustrated in Figure 7.

Relationship between ΔTPSim and task performance. To test for a behavioral relevance of the observed significant phase ΔTPSim clusters during WM maintenance, we investigated whether the power of low-frequency oscillations at times of significant sentence identity coding was related to task performance. Every wrongly reproduced sentence was defined as an error. Because nearly all error trials were expectedly different sentence trials, we could not investigate the direct relationship between phase coding in beta oscillations and behavior using ΔTPSim.

The number of electrodes coding sentence identity in phase ΔTPSim in the theta band and in the form of broadband gamma power ΔTPSim

Figure 4. Sentence identity coding (phase ΔTPSim) during verbal WM: group data of left perisylvian ECoG. Plots represent averages over electrodes and participants during the maintenance phase with trials time-locked on the stimulus onset (0 s, top panels), on the different stimulus offsets (0 s, middle panels), and on the individual speech onsets (0 s, bottom panels). Phase ΔTPSim based on the complex coefficient of phase-locking values of the time frequency-resolved data in the low-frequency range (4–48 Hz) are shown in each region. Significant clusters during WM maintenance are framed in white ($p < 0.01$, stepwise Bonferroni-corrected). Borders of clusters that are too close to stimulus offset or the speech onset to be interpreted as WM-related have been colored in black. Temporal information coding (phase ΔTPSim) was observed during WM in the pars triangularis of the IFG in all of the three analyses. Phase ΔTPSim in motor cortices was first observed after the stimulus was perceived. Phase ΔTPSim in auditory association cortices was primarily observed in WM processes related to stimulus onset, respectively, to speech onset. Time points on the *x* axes correspond to the middle of the 500 ms analysis window. Significant clusters at the borders of the maintenance phase were not interpreted further because these effects could have been generated by the listening or the speaking phase. S, Speaking.

was not large enough to permit testing a relationship with behavior. Because phase coding is likely modulated by the amplitude of the underlying oscillation (more robust phase coding with increasing power) (Wang, 2010), we tested whether power of low-frequency oscillations in significant clusters of the corresponding phase ΔTPSim during WM maintenance was related with correct sentence reproduction. The time-frequency clusters that were entered in this analysis are illustrated in Figure 4 (top panels): pars triangularis of the IFG (beta: centered at 2 s and 29.5 Hz; 2.15 s and 13 Hz), pars opercularis of the IFG (alpha: centered at 2.05 s and 9 Hz), mid STG (beta: centered at 1.7 s and 27 Hz), posterior STG (alpha: centered at 2.3 s and 8.5 Hz, beta: centered at 3 s and 15 Hz, low gamma: centered at 1.8 s and 46.5 Hz; 2.55 s and 45.5 Hz); Figure 4 (middle panels): pars triangularis of the IFG (beta: centered at 1.05 s and 14.5 Hz), premotor cortex (low gamma: centered at 0.9 s and 42 Hz), ventral motor cortex (beta: centered at 0.75 s 30 Hz), dorsal motor cortex (low gamma: centered at 0.45 s and 43.5 Hz; 0.55 s and 45.5 Hz); Figure 4 (bottom panels): pars triangularis of the IFG (alpha: centered at −1.25 s and 11 Hz, beta: centered at −0.7 s and 24.5 Hz; −0.2 s and 17 Hz, low gamma: centered at −0.15 s and 36 Hz), dorsal motor cortex (alpha: centered at −0.35 s and 9 Hz, beta: centered at −0.65 s and

29 Hz), mid STG (alpha: centered at −0.95 s and 10 Hz, low gamma: centered at −1.1 s and 46 Hz), posterior STG (beta: centered at −0.35 s and 22 Hz). Baseline-corrected power in these clusters was averaged separately in correct different sentence trials and in incorrect different sentence trials (on average, 54 correct vs 7 incorrect different sentence trials). We calculated the difference between the median power of low-frequency oscillations in correct and incorrect trials in each individual cluster.

Statistical analysis. To test for significant differences between same and different sentence trial ΔTPSim (phase and power), we created surrogate data to obtain a null distribution for the ΔTPSim values. In each patient, all same sentence trials were used, and the same number of different sentence trials was selected randomly. The trials were randomly split in two equal half-sets, and the surrogate TPSim was calculated within each half-set and averaged. Surrogate ΔTPSim between the two half-sets was calculated by subtracting the two surrogate TPSims. This procedure was repeated 1000 times. Because beta effects are more distributed within regions compared with more local gamma effects (Courtemanche et al., 2003; Howe et al., 2011), the electrodes' real and surrogate phase ΔTPSim were averaged separately within each region. The real ΔTPSim in each
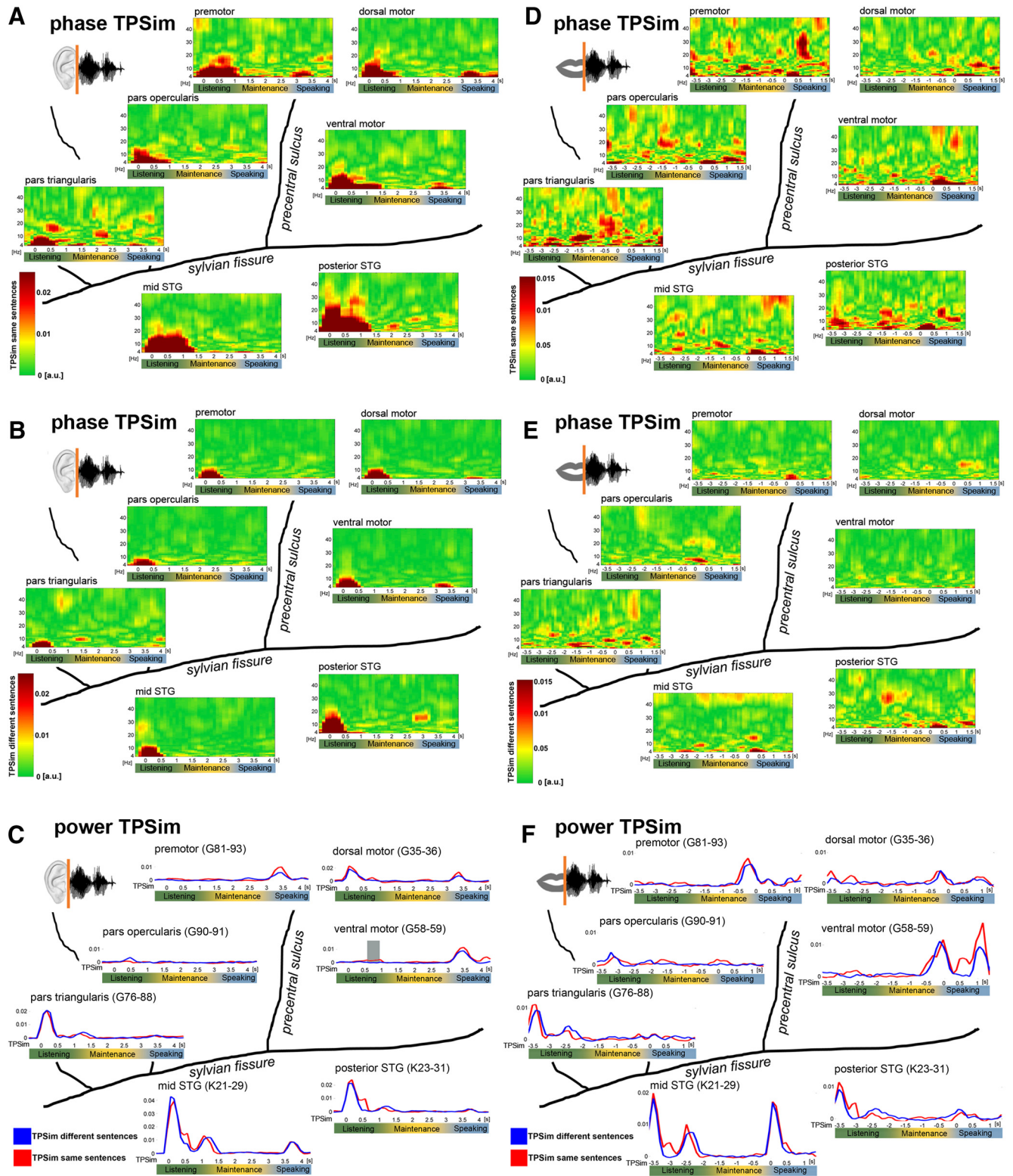
Figure 5. Input- and output-related phase and power TPSim. Time points on the *x* axes correspond to the center of the 500 ms analysis window. *A*, *B*, Input-related phase TPSim based on the complex coefficient phase-locking value of the time frequency-resolved data in the low-frequency range (4 – 48 Hz) of same (*A*) and different (*B*) sentence trials. Trials were time-locked on the auditory stimulus onset (0 s). Plots represent averages over electrodes and patients. *D*, *E*, Output-related phase TPSim based on the complex coefficient phase-locking value of the time frequency-resolved data in the low-frequency range (4 – 48 Hz) of same (*D*) and different (*E*) sentence trials. Trials were time-locked on individual speech onset (0 s). Plots represent averages over electrodes and patients. Input-related (*C*) and output-related (*F*) gamma power TPSim (70 – 170 Hz) of same (red) and different sentence trials (blue). *C*, Trials were time-locked on stimulus onset (0 s). *F*, Trials were time-locked on individual speech onset (0 s). Gray shaded area represents significant ΔTPSim time intervals (*p* = 0.0243, corrected, ventral motor cortex in *C*). Plots illustrate representative single electrodes in participant 3.
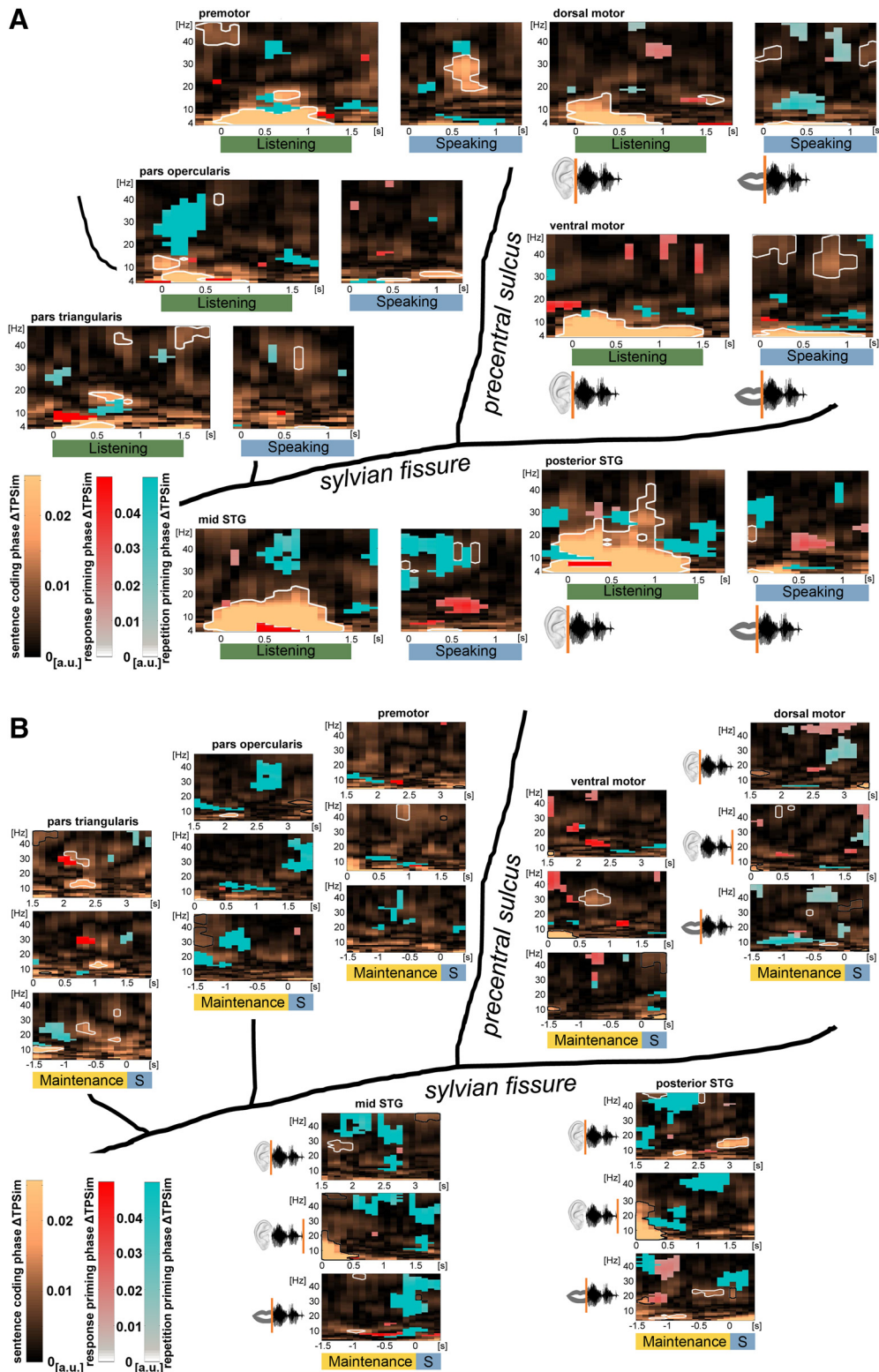
Figure 6. *A*, Response and repetition priming phase ΔTPSim during speech perception and production: group data of left perisylvian ECoG. Plots represent averages over electrodes and participants. For perception-related response and repetition priming phase ΔTPSim during listening (green), trials were time-locked on stimulus onset; whereas for the analysis of production-related response and repetition priming during speaking (blue), trials were time-locked on individual speech onsets. The analysis was performed on the complex coefficient phase-locking value of the time frequency-resolved data in the low-frequency range (4–48 Hz). Significant clusters (*p* < 0.01, stepwise Bonferroni-corrected) reflecting response priming (red) and repetition priming (blue) are overlaid onto the phase ΔTPSim analyses on sentence identity (related to Fig. 3). Time points on the *x* axes correspond to the middle of the 500 ms analysis window. *B*, Response and repetition priming phase ΔTPSim during verbal WM: group data of left perisylvian ECoG. Plots represent averages over electrodes and participants during WM maintenance with trials time-locked on the stimulus onset (0 s, top panels), on the different stimulus offsets (0 s, middle panels), and on the individual speech onsets (0 s, bottom panels). The analysis was performed on the complex coefficient phase-locking value of the time frequency-resolved data in the low-frequency range (4–48 Hz). Significant clusters (*p* < 0.01, stepwise Bonferroni-corrected) on response priming (red) and repetition priming (blue) are overlaid onto the phase ΔTPSim analyses, reflecting sentence identity coding during verbal WM (related to Fig. 4). Time points on the *x* axes correspond to the middle of the 500 ms analysis window. S, Speaking.
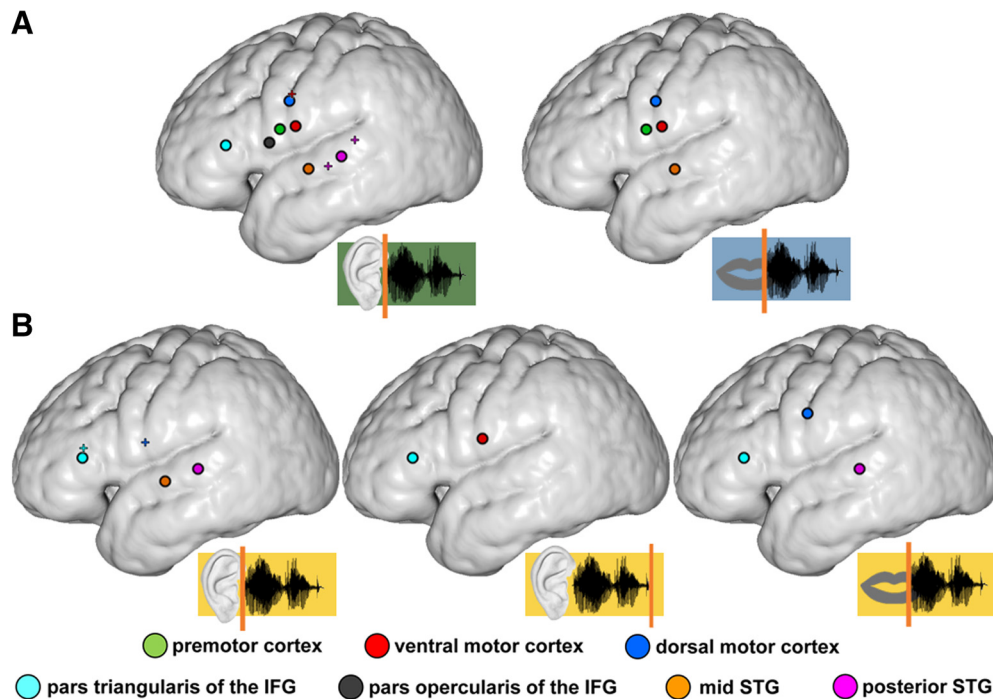
Figure 7. Summary figure of sentence identity coding in the beta phase ΔTPSim and the gamma power ΔTPSim. Black-rimmed circles represent regions that coded sentence identity in the phase of beta oscillations. Small colored crosses represent single electrodes that coded sentence identity in broadband gamma power fluctuations. ***A***, Sentence identity coding during listening (left) and speaking (right). ***B***, Sentence identity coding during WM maintenance with trials time-locked on the stimulus onset (left), on the different stimulus offsets (middle), and on the individual speech onsets (right). Whereas input-related sentence identity coding in the beta band was observed in the auditory association cortices and the IFG, beta phase coding in the motor cortex depended on the perception of the entire sentence. There is an overall paucity of electrodes coding sentence identity in the broadband gamma signal compared with coding in the beta band.

region was averaged separately over participants. Each participant provided 1 random of 1000 surrogate ΔTPSim for each region. The mean of those selected surrogate ΔTPSim was compared against the observed average ΔTPSim within each region. This comparison was performed 10,000 times for each time window and for each frequency bin for the frequency resolved phase ΔTPSim and for each time window for the gamma power envelope fluctuations (power ΔTPSim) with the only difference that gamma power ΔTPSim was calculated in single electrodes. If in 95% of the comparisons, the observed ΔTPSim was higher than the surrogate ΔTPSim, then these data points were considered above threshold. To exclude effects of response or repetition priming, the suprathreshold data points were masked exclusively with the significant clusters revealed by the ΔTPSim analysis on repetition effects (see above). The statistics of the repetition effects analyses were identical. Multiple comparison correction was performed on the masked suprathreshold data points, clustered in time, based on temporal adjacency. To this end, the resulting binary Boolean vector containing the suprathreshold time windows was permuted 10,000 times to create a reference cluster distribution. We tested whether the size of the observed clusters was larger than the size of the surrogate, permuted clusters. To focus the analyses on the strongest effects, this comparison was performed for the four largest clusters using a stepwise Bonferroni correction and correcting for the five separate analyses (largest real cluster >99% of the largest random cluster, second largest cluster >99.5% of the second largest random cluster etc.) (Waldhauser et al., 2015). The results of the gamma power ΔTPSim analysis were additionally Bonferroni-corrected for the number of electrodes.

Due to the temporal smearing of the 500 ms analysis window for TPSim calculations, temporal information coding was only considered specific for WM maintenance, when it was observed well away from the offset of the auditory sentence or the speech onset. Significant clusters around the offset of the auditory sentence or speech onset were not interpreted because they likely reflect components of evoked responses. Those clusters were framed in black in Figures 4 and 6.

Averaging phase ΔTPSim over electrodes may potentially induce spurious effects. To exclude phase ΔTPSim effects induced by averaging over

electrodes and participants, additional statistical analyses were performed as described above, yet, in single electrodes separately. The only difference was that the real-phase ΔTPSim in a given electrode was compared with the 1000 surrogate-phase ΔTPSim in this electrode. The resulting-phase ΔTPSim clusters were tested for significant sentence identity coding as described above. We then tested whether single electrodes in individual participants showed significant sentence identity coding in phase ΔTPSim in frequencies and at times at which the group phase ΔTPSim clusters based on averages within regions were significant. We report the median number of individual electrodes contributing to significant clusters in the group analysis.

Power of low-frequency oscillations in time-frequency clusters that showed significant phase ΔTPSim in the aforementioned group analysis was tested for a relationship with task performance. We tested whether the power of low-frequency oscillations differed between correct and incorrect trials. Non-normally distributed (Kolmogorov–Smirnov test) power differences between correct and incorrect different sentence trials were tested against zero using a Wilcoxon signed-rank test (alpha = 0.05).

To detect those regions that contributed most to the overall effect, the same analysis was performed for each region separately.

Code accessibility. The custom MATLAB and FieldTrip code is available upon request.

Results

Nine patients undergoing awake tumor surgery in the left perisylvian region repeated prerecorded three-word sentences following a visual Go cue after maintaining the sentence in WM for 1.5 s (Fig. 1A). Time-frequency analysis revealed that, in the auditory cortex, suppression of low frequencies and broadband gamma activity was expectedly stronger during listening than during speaking (Fig. 2). The motor cortex showed the opposite pattern. Suppression of low frequencies (from the theta to the beta range) persisted in primary motor cortex during WM maintenance, whereas the premotor cortex showed a relative WM-related high beta and low gamma power increase. The pars opercularis of

the IFG showed some suppression of low frequencies, whereas this effect was not observed in the pars triangularis where activity increased in a broad frequency range with an overall frequency drift toward the beta range during WM maintenance. The mid STG showed stronger beta power during WM maintenance compared with the posterior STG.

Behavioral results

Same sentences were not repeated correctly in 1.17% ($\pm$1.68%) of same sentence trials; 10% ($\pm$6%) of the different sentence trials were error trials. The average speech onset time was 590 ms ($\pm$140 ms). The reaction time in same sentence trials (565 ms $\pm$150 ms) did not differ significantly from reaction time in different sentence trials (602 ms $\pm$143 ms) ($p = 0.091$; two-sample t test).

Sentence identity coding during listening and speaking

The time-frequency resolved phase ΔTPSim in the phase of low-frequency oscillations from 4 to 48 Hz, based on phase TPSim in that frequency range (Fig. 5), was first computed with the trials time-locked relative to the auditory stimulus onset at the beginning of the listening period (Fig. 3, left panels). This identifies temporal information coding that is temporally related with online stimulus processing. As expected, phase ΔTPSim increased immediately after stimulus onset and remained significantly high ($p < 0.01$, stepwise Bonferroni–Holmes correction) during most of the listening period in all recorded brain regions. Phase ΔTPSim during listening was most prominent in auditory association cortex compared with motor or prefrontal cortex. Phase ΔTPSim ranged in frequencies from the theta to the low beta band (4–20 Hz) with only slight differences between regions. Additionally, phase ΔTPSim increased in the high beta (20–30 Hz) and low gamma band (30–40 Hz) in posterior STG, and in the low gamma band in the premotor cortex, the pars opercularis, and triangularis of the IFG (Fig. 3, left panels). In the dorsal motor cortex, phase ΔTPSim was found in the low beta band at the end of the listening period (Fig. 3, left panels).

Despite comparable acoustic envelopes between sentences, sentence identity was also coded in the broadband gamma envelope during listening, yet in only three electrodes in the motor and auditory association cortex (see Fig. 7A, left; see Fig. 5C and F for power TPSim of same and different sentence trials separately). This demonstrates that speech is coded not only in the form of the spatial distribution of broadband gamma activity over electrodes (Flinker et al., 2010; Mesgarani and Chang, 2012; Pasley et al., 2012; Lotte et al., 2015; Cheung et al., 2016) but also temporally within electrodes in the form of broadband gamma power modulations (Tang et al., 2017) and phase-modulation of low-frequency oscillations.

In the speaking trial phase (data time-locked to the individual speech onsets), significant phase ΔTPSim was found in all areas, including the motor and the temporal auditory association cortex (Fig. 3, right panels). Coding in the latter region was expected given the processing of the auditory feedback during speech production by the temporal cortex (Flinker et al., 2010). Yet, phase ΔTPSim was observed in relative higher frequencies compared with the listening phase because tracking in the lower frequencies was restricted to a narrow theta band (4–6 Hz) in all regions. Strong phase ΔTPSim in higher frequencies was present in the high beta and low gamma band in motor cortices, in the low gamma band in the pars triangularis of the IFG, and in the high beta and low gamma range in the auditory association cortex (Fig. 3, right panels). During speaking, broadband gamma power ΔTPSim did not survive correction for multiple comparisons in any recorded electrode.

Sentence identity coding during verbal WM

Central to our main question and hypothesis was whether neural activity in the maintenance phase was temporally structured in a way that permits identification of same versus different sentences in WM. When the analysis was performed with the trials time-locked on the auditory stimulus onset, traces of sentence identity encoding (phase ΔTPSim) during the maintenance period were found in the mid and posterior STG and the pars triangularis and opercularis of the IFG (Fig. 4, top panels). The different regions coded sentence identity in different frequency bands: alpha in the pars opercularis of the IFG and posterior STG, low beta in the posterior STG and pars triangularis, high beta in the pars triangularis and the mid STG, and low gamma in the posterior STG and the pars triangularis of the IFG (Fig. 4, top panels). As in this analysis, all trials were time-locked to the onset of the auditory stimulus in the listening phase, the aforementioned encoding patterns in WM were more related to early input-related processing. There is an absence of significant phase ΔTPSim in the motor cortices in this analysis.

Sentence identity coding that was temporally related with the offset of the auditory stimuli and consequently depended on the concluded perception of the entire sentence was found in the low beta band in the pars triangularis of the IFG, in the high beta band in the ventral motor cortex, and in the low gamma band in the dorsal and premotor cortex (Fig. 4, middle panels).

Output-related sentence-specific processes during WM maintenance were detected in auditory association cortices, dorsal motor cortex, and the pars triangularis of the IFG when trials were time-locked to the individual speech onsets (Fig. 4, bottom panels). Phase ΔTPSim was primarily found in the high beta and low gamma band. Additional sentence identity coding was observed in the theta/alpha band in the posterior STG, mid STG, the pars triangularis, and the dorsal motor cortex (Fig. 4, bottom panels).

Sentence identity coding, as identified in the group analysis of data averaged over electrodes and participants, was not an artifact of averaging or pooling because single electrode phase ΔTPSim analyses confirmed significant coding in individual electrodes (median number of significant electrodes per region: 5 electrodes with a range from 1 to 14 electrodes).

Short periods of broadband gamma power envelope fluctuations coded sentence identity during WM significantly only in two single electrodes in the pars triangularis of the IFG and the dorsal motor cortex in the analysis that was time-locked to auditory sentence onset (see Fig. 7B).

In sum, the phase of low-frequency oscillations coded sentence identity during WM more consistently than broadband gamma power fluctuations. Phase coding during WM was observed most frequently in the beta band (10 times) and less often in the low gamma (7 times) and alpha band (5 times), whereas, in contrast to the listening and speaking period, sentence identity coding in the theta band occurred only once during WM. During WM maintenance, input-related sentence identity coding in the beta band, the frequency range that showed the most prominent sentence identity coding, was first observed in the auditory association cortices and the pars triangularis of the IFG (see Fig. 7B, left). Sentence identity coding in the beta band that depended on the concluded perception of the entire phrase was additionally observed in the ventral motor cortex (see Fig. 7B, middle). Finally, output-related sentence identity coding in the beta band during WM maintenance was observed in the dorsal motor cortex, the pars triangularis of the IFG, and the posterior STG (see Fig. 7B, right).

We investigated further whether the amplitude of low-frequency oscillations was behaviorally relevant when oscillatory

phase coded sentence identity during WM maintenance. Only beta power was positively related with performance. Median beta power was higher in correct compared with incorrect trials in the time-frequency windows when beta phase ΔTPSim was significant (power difference between correct and incorrect trials 1.033 a.u., $p = 0.033$, Wilcoxon signed rank test), indicating less suppressed beta-band power values for correct than incorrect sentence reproduction compared with baseline. This effect was primarily driven by beta power in the pars triangularis of the IFG (5.447 a.u., $p = 0.0137$, Wilcoxon signed rank test) and in the mid STG (1.956 a.u., $p = 0.0312$, Wilcoxon signed rank test), whereas in the other regions, the relationship between beta power and behavior did not reach significance. Alpha or low gamma power did not differ significantly between correct and incorrect trials (alpha power difference: 1.6834 a.u., $p = 0.0619$; low gamma power difference: 0.0768 a.u., $p = 0.6552$, Wilcoxon signed rank test).

Priming effects

There was minimal overlay between time(-frequency) clusters of sentence identity coding and clusters representing effects of response or repetition priming. Response priming during listening occurred primarily in the theta and low gamma range (Fig. 6A, red clusters). During speaking, strong response priming was observed in the auditory association cortices in the beta range. During WM maintenance, response priming effects and sentence identity coding in the beta band occurred close by in the pars triangularis of the IFG (Fig. 6B, red clusters). Response priming effects in broadband gamma power ΔTPSim were observed in two electrodes in the ventral motor cortex during listening, but not during speaking. During WM maintenance, significant broadband gamma power ΔTPSim response priming occurred only in one electrode in the dorsal motor cortex and in one electrode in the pars triangularis of the IFG.

Repetition priming showed larger effects than response priming. Later epochs during the experiment showed more consistent phase coding, particularly in the pars opercularis of the IFG and the auditory association cortices during listening, speaking, and WM maintenance (Fig. 6, blue clusters). With the exception of the posterior STG during listening and the mid STG during speaking, repetition priming effects occurred at relative higher frequencies in the respective frequency bands compared with sentence identity coding. Repetition priming effects during WM in the phase of low-frequency oscillations were primarily observed in the beta and low gamma band. Repetition priming effects in the broadband gamma power ΔTPSim were observed in electrodes in all regions, except the dorsal motor cortex and the posterior STG during listening. During speaking, no electrodes showed effects of response or repetition priming in the power ΔTPSim analyses. During WM maintenance, two electrodes in the dorsal motor cortex and two electrodes in the IFG showed effects of repetition priming in broadband gamma power ΔTPSim.

Discussion

We have demonstrated that sentence identity during WM is consistently represented in the phase of low-frequency oscillations, particularly in the beta range, in the studied left frontal and temporal cortical areas. To our knowledge, this report is the first to demonstrate that sentence encoding in WM occurs in the phase of beta oscillations in left frontotemporal regions.

While sentence identity during WM was also decoded from alpha and low gamma oscillations, sentence identity coding in the beta band was most prominent and related with task performance. This suggests a behaviorally relevant, yet not exclusive, coding of sentence identity in the beta band. In a theoretical speech decoding model (Ghitza, 2011), endogenous beta oscillations in auditory cortex are phase-reset by a speech-parsing theta rhythm to map beta-cycle long speech segments to memory neurons that represent phonetic features. This suggests that the access to linguistic memory representations is temporally structured by beta oscillations. Indeed, repetitive transcranial magnetic stimulation over the left IFG with stimulation frequencies in the beta, but not alpha or theta range, interferes with verbal memory (Hanslmayr et al., 2014). The special role of beta oscillations for temporal information coding in memory is not restricted to the verbal domain, as beta oscillations facilitate also visual WM (Buschman and Miller, 2007; Siegel et al., 2009; Düzel et al., 2010; Staudigl et al., 2015). In visual WM, prefrontal action potentials are aligned to beta oscillations in the local field potential in such a way that the phase of the beta oscillation codes the order of sequentially memorized items ("phase separation") (Siegel et al., 2009). This points to the role of beta oscillations in sequencing and timing (Arnal, 2012; Fujioka et al., 2012). Sequencing is important during speech and particularly sentence processing. Congruently, when WM-related processing serves syntactic unification during sentence perception, EEG beta power increases compared with perception of word lists. This has so far only been demonstrated on the scalp level (Bastiaansen et al., 2010; for conflicting results, see Lam et al., 2016). The phase of beta oscillations could represent sequential order in verbal memory, which could represent a prerequisite for correct speech reproduction. We show here that beta power was higher in correct compared with incorrect trials at times when beta phase coded sentence identity in WM. During WM maintenance, beta power was less suppressed compared with the listening and speaking trial phases and even increased compared with silent baseline (Fig. 2). Increased amplitudes of beta oscillations have been associated with internal timing processes during sequencing (Gompf et al., 2017). We hypothesize that relatively higher amplitudes of beta oscillations during sentence encoding facilitate phase separation and thus enhance the memory representation contained therein.

Sentence identity was decoded both from the phase of low-frequency oscillations and from the amplitude fluctuations of the broadband gamma signal. Yet, coding was much sparser in gamma power compared with low-frequency phase. When data were aligned to individual speech onsets, broadband gamma power modulations tracked the succession of regional activation from the motor to the auditory association cortices (Fig. 5F). Yet, broadband gamma power fluctuations did not significantly code sentence identity during speaking. Phonemes embedded in syllables or words have previously been decoded from motor and auditory cortex broadband gamma signals (Flinker et al., 2010; Korzeniewska et al., 2011; Pei et al., 2011; Bouchard et al., 2013; Cogan et al., 2014). Our negative finding could either result from slight trial-to-trial variability in the reproduction of the same sentence, or suggests that the phase of low-frequency oscillations carries more information on sentence context than the amplitude modulations of the broadband gamma signal. In the following, we thus focus the discussion on phase coding.

Our results confirm that content is represented in short bouts of neural activity during WM (Lundqvist et al., 2018a,b; Miller et al., 2018). Sentence identity was coded in every studied brain region in the phase of low-frequency oscillations already during listening (Fig. 7A, left), which suggests that, at least in the context of a sentence reproduction paradigm, information on perceived

speech is directly disseminated throughout the left frontotemporal speech network. Yet, our results suggest that WM maintenance represents a dynamic process with parallel input- and output-related information processing in the left perisylvian region. Input- as well as output-related information coding was found in both the frontal and temporal cortex, notably including the motor and auditory association cortex. Input-related sentence identity coding in the phase of low-frequency oscillations showed a temporal relationship with both sentence onsets and offsets. In the motor cortices, coding in the phase of low-frequency oscillations had a stronger relationship with sentence offsets than onsets (Fig. 7B), suggesting that sentence identity coding in the phase of low-frequency oscillations in these cortical areas is more dependent on the integrality of the planned utterance.

The fact that information can be decoded from brain signals does not automatically imply that the brain actively uses this information (Bouton et al., 2018). Yet, lesion data indeed propose a modulatory role of frontal cortical areas in speech perception (Hickok and Poeppel, 2007; Murakami et al., 2015) and an important contribution of temporal cortical areas to speech production (Indefrey and Levelt, 2004; Tourville and Guenther, 2011; Hickok, 2012). This suggests that speech representation in the brain is less modular compared with previous proposals of theoretical speech models. WM-related processes were detected also in the auditory association and motor cortex (see also Cogan et al., 2017). This specifies theoretical WM models in such a way that it includes those cortical areas that upon lesions produce not only WM deficits but also nonamnestic symptoms. The mid STG and the pars triangularis of the IFG showed the most significant relationship between WM-related beta oscillations and task performance, suggesting an important role of these cortices in verbal WM. Indeed, sentence-level memory load effects have been observed in anterior parts of Broca's area that is part of the ventral speech processing stream (Bonhage et al., 2014). Both lesions in Broca's region and in the anterior aspects of the left superior temporal cortex produce verbal WM deficits (Busch et al., 2015).

Temporal information coding during speaking was restricted to a narrow theta band, the high beta band, and the low gamma range compared with the much broader phase ΔTPSim during listening. The same auditory sentence was always a repetition of the identical recording while participants may have reproduced the same sentence slightly differently from trial to trial. This confirms that theta and gamma oscillations adapt to subtle rhythmic irregularities that may arise from slightly different speech tempi during speech production (Luo and Poeppel, 2007; Ghitza and Greenberg, 2009; Giraud and Poeppel, 2012; Gross et al., 2013). Note the almost absence of sentence identity coding in the theta band during WM maintenance compared with the listening and speaking trial phases in our experiment. Neural theta oscillations are entrained by the perceived syllable rate during perception and production (Luo and Poeppel, 2007; Giraud and Poeppel, 2012; Behroozmand et al., 2015). In consequence, our results suggest that sentences are stored in WM in a more abstract form than the syllabic level. Beta oscillations more strongly reflect internal computational brain rhythms associated with top-down processing compared with stimulus-processing theta and low gamma oscillations (Giraud and Poeppel, 2012; Bressler and Richter, 2015; Lee et al., 2015). A recent WM model proposes both bottom-up and top-down signals carry content-specific information and top-down beta oscillations regulate bottom-up flow of sensory information in the gamma range during WM (Salazar et al., 2012; Lundqvist et al., 2018b; Miller et al., 2018). Interareal synchronization of beta oscillations has been suggested to underlie the top-down implementation of neural ensembles (Roelfsema et al., 1997; Bastos et al., 2012; Michalareas et al., 2016; Miller et al., 2018). Because sentence identity coding in beta and low gamma oscillations was observed in different regions at different times of WM maintenance, it is likely that multiple, hierarchically organized, brain regions assist in representing complex information in WM.

Priming effects confirm proposed functional asymmetries in bottom-up and top-down processing in the theta/gamma versus beta range. Response priming during listening occurred primarily in the theta and low gamma range. This suggests reproducing the same sentence in successive trials shapes phase coding in oscillations associated with bottom-up sensory processing. Repetition priming in contrast to response priming is more closely related to memory representations on a larger time scale and may therefore involve additional top-down but also bottom-up signals during WM maintenance. Indeed, repetition priming during WM was primarily observed in the beta and low gamma range.

In conclusion, broadband gamma power fluctuations and the phase of local low-frequency oscillations in the left temporal and frontal cortex, particularly in the beta band, represent sentences in WM. The lack of temporal information coding in the theta band suggests that the brain codes sentences in memory in a more abstract form than at the syllabic level. The fact that sentence identity is consistently coded in the phase of beta oscillations and that beta amplitude during WM maintenance correlates at the same time with performance confirms the role of these neural signals in top-down processing.

References

Alderson-Day B, Fernyhough C (2015) Inner speech: development, cognitive functions, phenomenology, and neurobiology. Psychol Bull 141:931–965.

Arnal LH (2012) Predicting "when" using the motor system's beta-band oscillations. Front Hum Neurosci 6:225.

Ashburner J (2007) A fast diffeomorphic image registration algorithm. Neuroimage 38:95–113.

Baddeley A (2003) Working memory and language: an overview. J Commun Dis 36:189–208.

Bastiaansen M, Magyari L, Hagoort P (2010) Syntactic unification operations are reflected in oscillatory dynamics during on-line sentence comprehension. J Cogn Neurosci 22:1333–1347.

Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ (2012) Canonical microcircuits for predictive coding. Neuron 76:695–711.

Behroozmand R, Ibrahim N, Korzyukov O, Robin DA, Larson CR (2015) Functional role of delta and theta band oscillations for auditory feedback processing during vocal pitch motor control. Front Neurosci 9:109.

Bonhage CE, Fiebach CJ, Bahlmann J, Mueller JL (2014) Brain signature of working memory for sentence structure: enriched encoding and facilitated maintenance. J Cogn Neurosci 26:1654–1671.

Bonhage CE, Meyer L, Gruber T, Friederici AD, Mueller JL (2017) Oscillatory EEG dynamics underlying automatic chunking during sentence processing. Neuroimage 152:647–657.

Bouchard KE, Mesgarani N, Johnson K, Chang EF (2013) Functional organization of human sensorimotor cortex for speech articulation. Nature 495:327–332.

Bouton S, Chambon V, Tyrand R, Guggisberg AG, Seeck M, Karkar S, van de Ville D, Giraud AL (2018) Focal versus distributed temporal cortex activity for speech sound category assignment. Proc Natl Acad Sci U S A 115:E1299–E1308.

Bressler SL, Richter CG (2015) Interareal oscillatory synchronization in top-down neocortical processing. Curr Opin Neurobiol 31:62–66.

Buchsbaum BR, D'Esposito M (2008) The search for the phonological store: from loop to convolution. J Cogn Neurosci 20:762–778.

Busch RM, Love TE, Jehi LE, Ferguson L, Yardi R, Najm I, Bingaman W, Gonzalez-Martinez J (2015) Effect of invasive EEG monitoring on cog-

nitive outcome after left temporal lobe epilepsy surgery. Neurology 85:1475–1481.

Buschman TJ, Miller EK (2007) Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. Science 315:1860–1862.

Buzsáki G (2006) Rhythms of the brain. Oxford: Oxford UP.

Cheung C, Hamiton LS, Johnson K, Chang EF (2016) The auditory representation of speech sounds in human motor cortex. eLife 5:133.

Cogan GB, Thesen T, Carlson C, Doyle W, Devinsky O, Pesaran B (2014) Sensory-motor transformations for speech occur bilaterally. Nature 507:94–98.

Cogan GB, Iyer A, Melloni L, Thesen T, Friedman D, Doyle W, Devinsky O, Pesaran B (2017) Manipulating stored phonological input during verbal working memory. Nat Neurosci 20:279–286.

Courtemanche R, Fujii N, Graybiel AM (2003) Synchronous, focally modulated β-band oscillations characterize local field potential activity in the striatum of awake behaving monkeys. J Neurosci 23:11741–11752.

Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis: I. Segmentation and surface reconstruction. Neuroimage 9:179–194.

Ding N, Simon JZ (2014) Cortical entrainment to continuous speech: functional roles and interpretations. Front Hum Neurosci 8:311.

Düzel E, Penny WD, Burgess N (2010) Brain oscillations and memory. Curr Opin Neurobiol 20:143–149.

Flinker A, Chang EF, Kirsch HE, Barbaro NM, Crone NE, Knight RT (2010) Single-trial speech suppression of auditory cortex activity in humans. J Neurosci 30:16643–16650.

Fujioka T, Trainor LJ, Large EW, Ross B (2012) Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. J Neurosci 32:1791–1802.

Ghitza O (2011) Linking speech perception and neurophysiology: speech decoding guided by cascaded oscillators locked to the input rhythm. Front Psychol 2:130.

Ghitza O, Greenberg S (2009) On the possible role of brain rhythms in speech perception: intelligibility of time-compressed speech with periodic and aperiodic insertions of silence. Phonetica 66:113–126.

Giraud AL, Poeppel D (2012) Cortical oscillations and speech processing: emerging computational principles and operations. Nat Neurosci 15:511–517.

Golumbic ZE, Ding N, Bickel S, Lakatos P, Schevon CA, McKhann GM, Goodman RR, Emerson R, Mehta AD, Simon JZ, Poeppel D, Schroeder CE (2013) Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party." Neuron 77:980–991.

Gompf F, Pflug A, Laufs H, Kell CA (2017) Non-linear relationship between BOLD activation and amplitude of beta oscillations in the supplementary motor area during rhythmic finger tapping and internal timing. Front Hum Neurosci 11:582.

Grabner G, Janke AL, Budge MM, Smith D, Pruessner J, Collins DL (2006) Symmetric atlasing and model based segmentation: an application to the hippocampus in older adults. In: Medical image computing and computer-assisted intervention: MICCAI 2006 (Hutchison D, Kanade T, Kittler J, Kleinberg JM, Mattern F, Mitchell JC, Naor M, Nierstrasz O, Pandu Rangan C, Steffen B, Sudan M, Terzopoulos D, Tygar D, Vardi MY, Weikum G, Larsen R, Nielsen M, Sporring J, eds), pp 58–66. Berlin: Springer.

Gross J, Hoogenboom N, Thut G, Schyns P, Panzeri S, Belin P, Garrod S (2013) Speech rhythms and multiplexed oscillatory sensory coding in the human brain. PLoS Biol 11:e1001752.

Hanslmayr S, Matuschek J, Fellner MC (2014) Entrainment of prefrontal beta oscillations induces an endogenous echo and impairs memory formation. Curr Biol 24:904–909.

Henson RN, Eckstein D, Waszak F, Frings C, Horner AJ (2014) Stimulus-response bindings in priming. Trends Cogn Sci 18:376–384.

Herff C, Heger D, de Pesters A, Telaar D, Brunner P, Schalk G, Schultz T (2015) Brain-to-text: decoding spoken phrases from phone representations in the brain. Front Neurosci 9:217.

Herman AB, Houde JF, Vinogradov S, Nagarajan SS (2013) Parsing the phonological loop: activation timing in the dorsal speech stream determines accuracy in speech reproduction. J Neurosci 33:5439–5453.

Hickok G (2012) Computational neuroanatomy of speech production. Nat Rev Neurosci 13:135–145.

Hickok G, Poeppel D (2000) Towards a functional neuroanatomy of speech perception. Trends Cogn Sci 4:131–138.

Hickok G, Poeppel D (2004) Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 92:67–99.

Hickok G, Poeppel D (2007) The cortical organization of speech processing. Nat Rev Neurosci 8:393–402.

Howe MW, Atallah HE, McCool A, Gibson DJ, Graybiel AM (2011) Habit learning is associated with major shifts in frequencies of oscillatory activity and synchronized spike firing in striatum. Proc Natl Acad Sci U S A 108:16801–16806.

Indefrey P, Levelt WJ (2004) The spatial and temporal signatures of word production components. Cognition 92:101–144.

Jacquemot C, Scott SK (2006) What is the relationship between phonological short-term memory and speech processing? Trends Cogn Sci 10:480–486.

Korzeniewska A, Franaszczuk PJ, Crainiceanu CM, Kus R, Crone NE (2011) Dynamics of large-scale cortical interactions at high gamma frequencies during word production: event related causality (ERC) analysis of human electrocorticography (ECoG). Neuroimage 56:2218–2237.

Kriegeskorte N, Mur M, Bandettini P (2008) Representational similarity analysis: connecting the branches of systems neuroscience. Front Syst Neurosci 2:4.

Kubanek J, Brunner P, Gunduz A, Poeppel D, Schalk G (2013) The tracking of speech envelope in the human cortex. PLoS One 8:e53398.

Lachaux JP, Rodriguez E, Martinerie J, Varela FJ (1999) Measuring phase synchrony in brain signals. Hum Brain Mapp 8:194–208.

Lam NH, Schoffelen JM, Uddén J, Hultén A, Hagoort P (2016) Neural activity during sentence processing as reflected in theta, alpha, beta, and gamma oscillations. Neuroimage 15:43–54.

Lee JH, Whittington MA, Kopell NJ (2015) Potential mechanisms underlying intercortical signal regulation via cholinergic neuromodulators. J Neurosci 35:15000–15014.

Lotte F, Brumberg JS, Brunner P, Gunduz A, Ritaccio AL, Guan C, Schalk G (2015) Electrocorticographic representations of segmental features in continuous speech. Front Hum Neurosci 9:97.

Lundqvist M, Herman P, Miller EK (2018a) Working memory. delay activity, Yes! Persistent activity? Maybe not. J Neurosci 38:7013–7019.

Lundqvist M, Herman P, Warden MR, Brincat SL, Miller EK (2018b) Gamma and beta bursts during working memory readout suggest roles in its volitional control. Nat Commun 9:394.

Luo H, Poeppel D (2007) Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. Neuron 54:1001–1010.

Mercier MR, Bickel S, Megevand P, Groppe DM, Schroeder CE, Mehta AD, Lado FA (2017) Evaluation of cortical local field potential diffusion in stereotactic electro-encephalography recordings: a glimpse on white matter signal. Neuroimage 147:219–232.

Mesgarani N, Chang EF (2012) Selective cortical representation of attended speaker in multi-talker speech perception. Nature 485:233–236.

Mesgarani N, Cheung C, Johnson K, Chang EF (2014) Phonetic feature encoding in human superior temporal gyrus. Science 343:1006–1010.

Michalareas G, Vezoli J, van Pelt S, Schoffelen JM, Kennedy H, Fries P (2016) Alpha-beta and gamma rhythms subserve feedback and feedforward influences among human visual cortical areas. Neuron 89:384–397.

Michelmann S, Bowman H, Hanslmayr S (2016) The temporal signature of memories: identification of a general mechanism for dynamic memory replay in humans. PLoS Biol 14:e1002528.

Miller EK, Lundqvist M, Bastos AM (2018) Working memory 2.0. Neuron 100:463–475.

Murakami T, Kell CA, Restle J, Ugawa Y, Ziemann U (2015) Left dorsal speech stream components and their contribution to phonological processing. J Neurosci 35:1411–1422.

Nelson MJ, El Karoui I, Giber K, Yang X, Cohen L, Koopman H, Cash SS, Naccache L, Hale JT, Pallier C, Dehaene S (2017) Neurophysiological dynamics of phrase-structure building during sentence processing. Proc Natl Acad Sci U S A 114:E3669–E3678.

Nourski KV, Reale RA, Oya H, Kawasaki H, Kovach CK, Chen H, Howard MA 3rd, Brugge JF (2009) Temporal envelope of time-compressed speech represented in the human auditory cortex. J Neurosci 29:15564–15574.

Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput Intell Neurosci 2011:156869.

Pasley BN, David SV, Mesgarani N, Flinker A, Shamma SA, Crone NE, Knight

RT, Chang EF (2012) Reconstructing speech from human auditory cortex. PLoS Biol 10:e1001251.

Pei X, Leuthardt EC, Gaona CM, Brunner P, Wolpaw JR, Schalk G (2011) Spatiotemporal dynamics of electrocorticographic high gamma activity during overt and covert word repetition. Neuroimage 54:2960–2972.

Peña M, Melloni L (2012) Brain oscillations during spoken sentence processing. J Cogn Neurosci 24:1149–1164.

Perrone-Bertolotti M, Rapin L, Lachaux JP, Baciu M, Loevenbruck H (2014) What is that little voice inside my head? Inner speech phenomenology, its role in cognitive performance, and its relation to self-monitoring. Behav Brain Res 261:220–239.

Potter MC, Lombardi L (1990) Regeneration in the short-term recall of sentences. J Mem Lang 29:633.

Potter MC, Lombardi L (1998) Syntactic priming in immediate recall of sentences. J Mem Lang 38:265–282.

Rimmele JM, Zion Golumbic E, Schröger E, Poeppel D (2015) The effects of selective attention and speech acoustics on neural speech-tracking in a multi-talker scene. Cortex 68:144–154.

Roelfsema PR, Engel AK, König P, Singer W (1997) Visuomotor integration is associated with zero time-lag synchronization among cortical areas. Nature 385:157–161.

Salazar RF, Dotson NM, Bressler SL, Gray CM (2012) Content-specific fronto-parietal synchronization during visual working memory. Science 338:1097–1100.

Siegel M, Warden MR, Miller EK (2009) Phase-dependent neuronal coding of objects in short-term memory. Proc Natl Acad Sci U S A 106: 21341–21346.

Staudigl T, Vollmar C, Noachtar S, Hanslmayr S (2015) Temporal-pattern similarity analysis reveals the beneficial and detrimental effects of context reinstatement on human memory. J Neurosci 35:5373–5384.

Tang C, Hamilton LS, Chang EF (2017) Intonational speech prosody encoding in the human auditory cortex. Science 357:797–801.

Tourville JA, Guenther FH (2011) The DIVA model: a neural theory of speech acquisition and production. Lang Cogn Process 26:952–981.

Waldhauser GT, Bäuml KH, Hanslmayr S (2015) Brain oscillations mediate successful suppression of unwanted memories. Cereb Cortex 25:4180–4190.

Wang XJ (2010) Neurophysiological and computational principles of cortical rhythms in cognition. Physiol Rev 90:1195–1268.

Yaffe RB, Kerr MS, Damera S, Sarma SV, Inati SK, Zaghloul KA (2014) Reinstatement of distributed cortical oscillations occurs with precise spatiotemporal dynamics during successful memory retrieval. Proc Natl Acad Sci U S A 111:18727–18732.