

Sleep Disrupts High-Level Speech Parsing Despite Significant Basic Auditory Processing

Shiri Makov,^{1,3} Omer Sharon,¹ Nai Ding,² Michal Ben-Shachar,^{3,4} Yuval Nir,^{1,5,6*} and Elana Zion Golumbic^{3*}

¹Sagol School of Neuroscience, Tel Aviv University, Tel Aviv 69978, Israel, ²College of Biomedical Engineering and Instrument Sciences, Zhejiang University, Hangzhou 310027, China, ³The Gonda Center for Multidisciplinary Brain Research, and ⁴Department of English Literature and Linguistics, Bar Ilan University, Ramat Gan, 5290002, Israel, ⁵Department of Physiology and Pharmacology, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel, and ⁶Functional Neurophysiology and Sleep Research Laboratory, Tel-Aviv Sourasky Medical Center, 64239 Tel-Aviv, Israel

The extent to which the sleeping brain processes sensory information remains unclear. This is particularly true for continuous and complex stimuli such as speech, in which information is organized into hierarchically embedded structures. Recently, novel metrics for assessing the neural representation of continuous speech have been developed using noninvasive brain recordings that have thus far only been tested during wakefulness. Here we investigated, for the first time, the sleeping brain's capacity to process continuous speech at different hierarchical levels using a newly developed Concurrent Hierarchical Tracking (CHT) approach that allows monitoring the neural representation and processing-depth of continuous speech online. Speech sequences were compiled with syllables, words, phrases, and sentences occurring at fixed time intervals such that different linguistic levels correspond to distinct frequencies. This enabled us to distinguish their neural signatures in brain activity. We compared the neural tracking of intelligible versus unintelligible (scrambled and foreign) speech across states of wakefulness and sleep using high-density EEG in humans. We found that neural tracking of stimulus acoustics was comparable across wakefulness and sleep and similar across all conditions regardless of speech intelligibility. In contrast, neural tracking of higher-order linguistic constructs (words, phrases, and sentences) was only observed for intelligible speech during wakefulness and could not be detected at all during nonrapid eye movement or rapid eye movement sleep. These results suggest that, whereas low-level auditory processing is relatively preserved during sleep, higher-level hierarchical linguistic parsing is severely disrupted, thereby revealing the capacity and limits of language processing during sleep.

Key words: attention; entrainment; sleep; speech processing

Significance Statement

Despite the persistence of some sensory processing during sleep, it is unclear whether high-level cognitive processes such as speech parsing are also preserved. We used a novel approach for studying the depth of speech processing across wakefulness and sleep while tracking neuronal activity with EEG. We found that responses to the auditory sound stream remained intact; however, the sleeping brain did not show signs of hierarchical parsing of the continuous stream of syllables into words, phrases, and sentences. The results suggest that sleep imposes a functional barrier between basic sensory processing and high-level cognitive processing. This paradigm also holds promise for studying residual cognitive abilities in a wide array of unresponsive states.

Introduction

Sleep is defined as a reversible state in which external stimuli rarely affect perception or elicit meaningful behavioral responses

(Nir and Tononi, 2010). Despite such disconnection, it is clear that some discriminative sensory processing persists during sleep. Recent studies, particularly in the auditory domain, found preserved activation of primary sensory cortices during sleep

Received Jan. 19, 2017; revised April 27, 2017; accepted May 3, 2017.

Author contributions: S.M., N.D., M.B.-S., Y.N., and E.Z.G. designed research; S.M. and O.S. performed research; N.D., Y.N., and E.Z.G. contributed unpublished reagents/analytic tools; S.M. and O.S. analyzed data; S.M., Y.N., and E.Z.G. wrote the paper.

This work was supported by the I-CORE Program of the Planning and Budgeting Committee and the Israel Science Foundation (Y.N. and E.Z.G.), an FP7 Marie Curie Career Integration Grant (Y.N. and E.Z.G.), the Israel Science Foundation (Grant 1326/15 to Y.N.), the Binational Science Foundation (BSF) (Grant 2015385 to E.Z.G.), and the Adelis Foundation (Y.N.). We thank Talma Hendler for continuing support at the Tel Aviv Sourasky Medical Center, Shani Shalgi for help setting up the EEG sleep laboratory, Shlomit Beker for assistance in setting up the experiment,

Noam Amir for advising on acoustic aspects of stimulus preparation, Netta Neeman for assistance with data acquisition, Noa Bar-Ilan Regev for administrative help, and Yaniv Sela and laboratory members for suggestions.

The authors declare no competing financial interests.

*Y.N. and E.Z.G. are co-senior authors.

Correspondence should be addressed to either of the following: Elana Zion-Golumbic, The Gonda Center for Multidisciplinary Brain Research, Bar Ilan University, Ramat Gan 5290002, Israel, E-mail: elana.zion-golumbic@biu.ac.il; or Yuval Nir, Sagol School of Neuroscience, Tel Aviv University, Tel Aviv 69978, Israel, E-mail: ymir@post.tau.ac.il.

DOI:10.1523/JNEUROSCI.0168-17.2017

Copyright © 2017 the authors 0270-6474/17/377772-10\$15.00/0

(Peña et al., 1999; Portas et al., 2000; Issa and Wang, 2008; Andrillon et al., 2015). Accordingly, single-neuron responses to both simple (click or tone) and complex (vocalization) stimuli are comparable to those in wakefulness, including frequency tuning curves and stimulus-specific adaptation to deviant sounds (Issa and Wang, 2008; Nir et al., 2015). Therefore, external stimuli give rise to robust sensory representations in the sleeping brain, yet the extent of their processing remains unclear.

The uncertainty regarding the depth of processing is enhanced with respect to complex stimuli such as speech, which require both sensory and high-order processing, engaging multiple brain areas beyond auditory cortex (Peelle et al., 2010; Wylie and Regner, 2014; Friederici and Singer, 2015). To date, there is limited and inconsistent evidence regarding the level of processing that speech undergoes during sleep. Several studies have reported that, during sleep, activity in high-order language-related regions such as the superior temporal gyrus (STG), temporal-parietal junction (TPJ), and the inferior frontal gyrus (IFG) is robustly attenuated or absent altogether and brain responses to regular speech and to meaningless control conditions are similar (Portas et al., 2000; Dehaene-Lambertz et al., 2002; Wilf et al., 2016). However, other evidence suggests that some level of semantic information is extracted even during sleep. For example, behaviorally relevant stimuli such as one's name lead to more frequent awakenings (Oswald et al., 1960; Langford et al., 1974; McDonald et al., 1975) and induce a spread of cortical activation (Pratt et al., 1999; Portas et al., 2000; Blume et al., 2017). In addition, studies measuring event-related potentials (ERPs) during sleep have demonstrated an N400 response to lexical-level semantic violations (Brualla et al., 1998; Bastuji et al., 2002; Ibáñez et al., 2006) and residual neural signatures indicating semantic categorization of words (Kouider et al., 2014; Andrillon et al., 2016). Notably, semantic-level responses during sleep are often considerably attenuated and delayed in time compared with those in wakefulness (Brualla et al., 1998; Perrin et al., 2002; Andrillon et al., 2016), raising important questions as to the nature of residual high-level language processing during sleep.

A main challenge for addressing this question is how to assess speech processing depth in lieu of behavioral metrics. Previous studies have resorted to measuring ERPs. However, during non-rapid eye movement (NREM) sleep, repetitive presentation of brief isolated stimuli often elicits a large stereotypical response known as a "K complex" (Colrain, 2005; Halász, 2016) that masks the precise neuronal dynamics and limits data interpretation. Indeed, ERPs recorded during NREM sleep often differ substantially in their time course and morphology from those observed during wakefulness (Colrain and Campbell, 2007), making it difficult to assess their functional significance.

Here, we used the newly developed concurrent hierarchical tracking (CHT) approach (Ding et al., 2016) to assess the depth of speech processing during sleep. In CHT, stimuli are structured in a manner that allows distinguishing neural responses to different levels of linguistic analysis of continuous speech. Specifically, speech sequences are compiled such that different linguistic levels (syllables, words, phrases, and sentences) correspond to distinct frequencies, allowing us to distinguish their neural signatures in brain activity while refraining from presentation of abrupt stimuli. We used this method to test whether hierarchical parsing of intelligible speech is preserved or disrupted during sleep. This approach allows examination of the capacity and limits of processing during sleep and may afford insights regarding the depth of speech processing that occurs without attention more generally. We hypothesized that, during sleep, the neural representa-

tion of higher linguistic levels will be greatly diminished, as has been shown during wakefulness for unintelligible or unattended speech (Zion Golumbic et al., 2013; Ding et al., 2016).

Materials and Methods

Participants. Full-night sleep recordings were performed in 29 native Hebrew speakers (16 females, mean age 28.7 ± 3.6 years, range 22–38) who reported to be healthy and without any history of neuropsychiatric or sleep disorders. The study was approved by the Medical Institutional Review Board at the Tel Aviv Sourasky Medical Center. All participants provided their written consent for participation. Participants underwent an interview determining their sleep habits and their propensity to fall asleep in noisy environments. Eight participants (6 females) were excluded from analysis for either technical issues ($n = 3$) or lack of sufficient data ($n = 5$ participants who kept falling asleep during the wake sessions or experienced difficulties sleeping during the night). Twenty-one participants were included in the wake analysis (10 females, mean age 28.2 ± 4.0), of which 17 were included in the NREM sleep analysis (6 females, mean age 28.8 ± 4.2) and 13 were included in the REM sleep analysis (5 females, mean age 28.6 ± 4.3).

Hebrew materials. A bank of individually recorded Hebrew syllables was used to create a set of intelligible speech stimuli, as well as unintelligible (scrambled) control stimuli. Single syllables were uttered in random order by a human male voice and recorded at 44,100 Hz. The sound intensity was manually normalized using Audacity software (version 2.0.5). To prevent biasing of speech perception by prosody, prosodic cues were removed via pitch normalization using Praat (version 5.4.04) (Boersma and Weenink, 2016). The length of individual syllables (original mean duration 243.6 ± 64.3 ms, range 168–397 ms) was adjusted to precisely 250 ms by truncation or silence padding at the end. In case of truncation, a fading out effect was applied on the last 25 ms. Finally, syllables were concatenated using custom-written scripts in MATLAB (The MathWorks) to form 25 intelligible Hebrew sentences and 25 unintelligible (scrambled) sequences. All intelligible sentences were constructed to form hierarchical linguistic structures as follows: every two syllables formed a 500-ms-long word, every two words formed a 1000-ms-long phrase, and every two phrases formed a 2000-ms-long sentence (Fig. 1A). Because syllables were grouped hierarchically into linguistic constituents with no additional acoustic gaps inserted between them (Fig. 1B), the linguistic structures appeared at fixed periodicities throughout the stimuli (syllables at 4 Hz, words at 2 Hz, phrases at 1 Hz, and sentences at 0.5 Hz). Sentences did not include rhymes, passive form of verbs, or arousing semantic content. The intelligibility of Hebrew materials was verified in a pilot study demonstrating that all Hebrew sentences could be fully repeated after a mean of 1.28 presentations and most ($78.7 \pm 15.8\%$) could be fully repeated by participants after a single presentation. The same Hebrew syllables were scrambled to compose 25 unintelligible pseudo-sentences. Scrambling was performed by shuffling syllables across sentences while maintaining their original position within the sequence. Therefore, syllables that tend to occur at the beginning/end of words in natural language retain this position in the control condition. We also verified that the scrambling procedure did not produce any real Hebrew words by chance.

Chinese materials. We used sentences in Chinese as an unintelligible speech control condition. The Chinese materials were constructed and used by Ding et al. (2015) in a similar manner as the Hebrew materials and were composed of sequences of individual syllables lasting 250 ms uttered by a computerized male voice (see Stimuli I in Ding et al., 2016). The Chinese materials were also compiled such that higher linguistic structures appeared at fixed frequencies. However, because the Chinese sentences were composed of monosyllabic words, the syllabic rate of 4 Hz also represented the word rate. Therefore, only two additional linguistic levels were formed: phrases at 2 Hz and sentences at 1 Hz. The fact that the Hebrew materials included an additional linguistic level should not have interfered with our experimental design because none of the participants understood Chinese and it was only used as an unintelligible speech control condition.

Stimulus presentation. A trial lasted 12 s and was composed of 6 concatenated 2-s-long sentences of the same speech type (intelligible, scram-

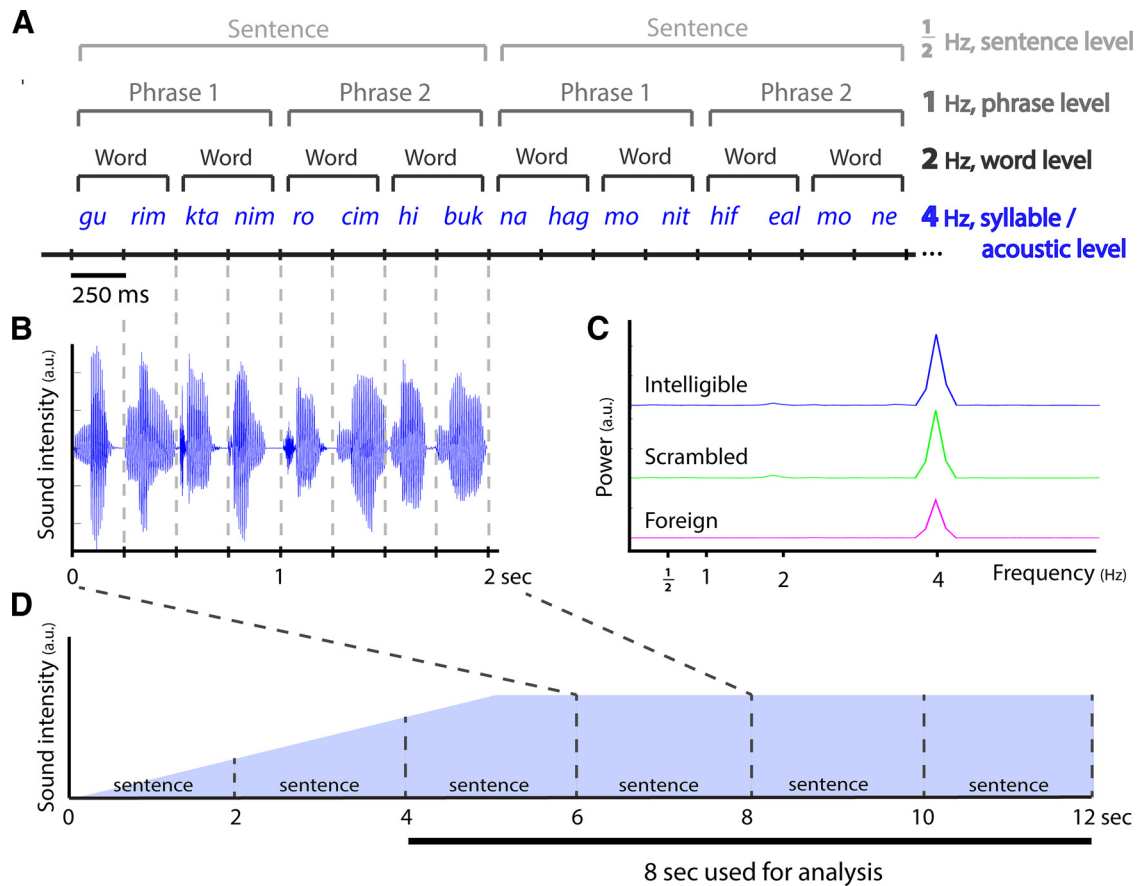


Figure 1. Stimuli with hierarchical linguistic structures at fixed frequencies used in the CHT paradigm. **A**, Example of intelligible (Hebrew) speech composed of 250 ms syllables in which 4 levels of information are differentiated based on their rate: acoustic/syllabic, word, phrasal, and sentential rates (at 4, 2, 1, and 0.5 Hz, respectively). Translation of the Hebrew sentences in the example: “Small puppies want a hug” and “A taxi driver turned on the meter.” Control stimuli were unintelligible syllable sequences with similar syllabic rate of 4 Hz forming either “scrambled” Hebrew speech or speech in a foreign language (Chinese). **B**, Representative sound wave of a single sentence (2 s). Sound intensity fluctuates at the rate of 4 Hz. **C**, Power spectrum of mean soundwave envelope across 50 trials per condition (blue, intelligible; green, scrambled; pink, foreign). Note that, across all conditions, the soundwave envelope exhibits a strong peak at the acoustic rate (4 Hz) with no significant power at lower frequencies corresponding to higher-order linguistic structures. **D**, In each 12 s trial, six sentences were concatenated without gaps. Trials began with a gradual increase in sound intensity to prevent sleep interference.

bled, or foreign speech; Fig. 1D). In the first 5 s of each trial, sound intensity increased from zero to full intensity to prevent abrupt onsets. To assure that the sound intensity of the stimuli fluctuated at the syllabic rate of 4 Hz, the power spectrum of the mean sound intensity in each condition was calculated across 50 trials for each condition (Fig. 1C). Intertrial intervals were distributed pseudorandomly between 1.5 and 4 s to avoid expectation effects. In addition to the three speech conditions, we presented a baseline sham condition in which no sound was presented. Typically, during a full experiment, we presented 200–500 trials (mean 406 ± 87) for each condition (intelligible, scrambled, foreign, and sham). Trials were not presented in a discrete block design, but rather in a continuous manner throughout the experiment. Approximately 20% of the trials were presented during wakefulness, 10% during REM sleep, 65% during NREM sleep, and an additional 5% were presented during state transitions (excluded). Only the final 8 s of each trial were used for analysis to further avoid onset responses (Fig. 1D).

Experimental design. The experiment took place in an acoustically attenuated and electrically shielded sleep laboratory. Auditory stimulation was delivered through loudspeakers and sound intensity was adjusted to a convenient level that remained constant throughout the night (see “Auditory stimulation” section below). Hebrew sentences were confirmed to be intelligible before the commencement of the experiment (see “Intelligibility test” section below). Experiments included stimulation during wakefulness in the evening, stimulation during overnight sleep, and another block of stimulation in the morning after spontaneous awakening (see Fig. 3A). The room was dark throughout the entire experiment. During the wakefulness sessions, participants sat on a chair

rather than lying in bed to ensure that they stayed awake and so that we could observe if any signs of falling asleep (e.g., rolling eye movements, EEG slowing, or appearance of sleep spindles or K-complexes); if any of these occurred, participants were woken up immediately (3.85 ± 3.21 , range 0–11 awakenings per participant). To minimize differences between wake and sleep sessions, participants were instructed to keep their eyes closed and were not required to perform any explicit task. After the evening wakefulness session, participants were allowed to fall asleep at their own convenience. Auditory stimulation was paused manually whenever awakening or movement was detected and resumed shortly after detection of unequivocal sleep activity in the EEG.

Auditory stimulation. Auditory stimulation was delivered through speakers situated on both sides of the bed (during sleep) or on both sides of the chair (during wakefulness). Sound intensity was individually adjusted to a relatively low level (range 42.1–45.9 dB SPL) that allowed speech comprehension. Importantly, sound intensity was adjusted before the first experimental session and remained constant throughout all conditions for each participant.

Intelligibility test. Participants performed an intelligibility test to verify that they understood the materials used in the intelligible speech condition. All intelligible and scrambled sentences were presented in random order and participants were asked to report whether they contained meaningful speech and were then asked to repeat them.

Data acquisition. High-density EEG was recorded continuously using a 256-channel hydrogel geodesic sensor net with passive electrodes (Electrical Geodesics). Each carbon fiber electrode consists of a silver chloride carbon fiber pellet, a lead wire, and a gold-plated pin and was injected

with conductive gel (Electro-Cap International). Signals were referenced to Cz, amplified via an AC-coupled high-input impedance amplifier (NetAmps 300; Electrical Geodesics), and digitized at 1000 Hz. Electrode impedance in all sensors was verified to be <50 k Ω before starting the recording.

Sleep scoring. Sleep scoring in 30 s epochs was performed manually according to established guidelines of the American Academy of Sleep Medicine (Iber et al., 2007) based on EEG, EOG, EMG, and video. To this end, EEG data from F3/F4, C3/C4, and O1/O2 were referenced to the contralateral mastoid and two EOG channels were referenced to the other mastoid. Scoring channels were visualized along with synchronized EMG in 30 s epochs. Sleep scoring was further verified by inspecting the time–frequency representation (spectrogram) of the Pz electrode (not involved in scoring process) superimposed with the hypnogram (as in Fig. 3*A,B*). Each epoch was categorized as N1/N2/N3/REM sleep or wakefulness. N1 sleep stage epochs ($13.4 \pm 1.4\%$ of sleep time, range 7.6–30.1%) were excluded from further analysis to avoid uncertainty regarding precise sleep onset.

EEG preprocessing. Preprocessing was performed in MATLAB (The MathWorks) using the FieldTrip toolbox (Oostenveld et al., 2011) and custom-written scripts. Data were segmented (–2 to +14 s) around stimulus onset, down-sampled to 250 Hz, rereferenced to the average signal of the mastoids, and linearly detrended. Bad electrodes (<14% in all participants) were identified as those in which variance and maximal absolute value constituted outliers relative to adjacent electrodes upon visual inspection per participant and per state and were replaced with the weighted average of their neighbors using a linear, distance-weighted interpolation. Outlier trials ($15.3 \pm 0.3\%$) were identified manually by visual inspection and were discarded from subsequent analysis. In the N2 dataset, we additionally excluded trials containing K-complex events. K-complexes were detected automatically as those trials in which the raw EEG amplitude was both higher than +40 μ V and lower than –40 μ V within a 2.4 s window, sliding at a resolution of 40 ms. Independent component analysis (ICA) was used for removal of eye movement and heartbeat traces separately for each state. After cleaning, we randomly selected the same number of trials from all conditions (intelligible, scrambled, foreign and sham) for each subject separately. Participants with fewer than 30 clean trials in any single condition were discarded from further analysis.

Data analysis during wakefulness. Data were analyzed by two complementary approaches. First, intertrial phase coherence (ITPC) was calculated as follows: the fast Fourier transform (FFT) was calculated separately for each trial with 0.125 Hz resolution. Next, the phase component at each frequency was used to calculate the ITPC, which is the sum (absolute value) of the phases across trials, as follows:

$$ITPC = \frac{1}{N} \sum_{k=1}^N e^{i\phi_k}$$

Note that ITPC represents phase consistency across trials, which is the inverse of response variability across trials (Berens, 2009).

Second, evoked power spectrum analysis was performed as follows: We averaged the clean preprocessed trials for each condition separately and computed the power spectrum of the average using FFT with 0.125 Hz resolution. We normalized the power at each frequency by subtracting the mean power level at adjacent frequencies within ± 0.125 Hz (Nozaradan et al., 2011).

Comparison of sleep and wakefulness. To reduce the effects of widespread slow waves during sleep (see Fig. 3*C*) that could preclude analysis of low frequencies of interest (0.5, 1 and 2 Hz), we applied a spatial current source density transformation (CSD, also known as surface Laplacian; Kayser and Tenke, 2015) to the preprocessed EEG data. Indeed, EEG spectral power after CSD transformation revealed comparable energy at slow (<4 Hz) frequencies across wakefulness, N2 sleep, and REM sleep (see Fig. 3*D*), attesting to the utility of this procedure in minimizing the potential effect of slow ongoing sleep activities. However, N3 sleep data were still dominated by robust slow-wave activity (see Fig. 3*D*) and therefore were excluded from analysis. Furthermore, for comparing wakefulness and sleep states, we focused on ITPC analysis because

phase consistency is less affected than power by ongoing slow activities. Importantly, identical procedures were applied across all states of sleep and wakefulness (see Fig. 4). Finally, when comparing the results in each sleep stage with wakefulness, we used the same participants for each comparison and randomly selected an equal number of trials across states to ensure similar statistical power. Topographical distribution of ITPC values in Figure 4 was calculated for intelligible speech after normalizing ITPC at each frequency by subtracting the mean ITPC at adjacent frequencies within ± 0.125 Hz.

SNR estimation. SNR (used as a covariate in subsequent ANCOVA tests) was quantified in each state and subject separately as follows:

$$SNR = \frac{SD \text{ of ERP, Intelligible condition}}{SD \text{ of ERP, Sham condition}}$$

Statistical analysis. Statistical analysis was performed using custom-written MATLAB scripts (The Mathworks) and SPSS software (version 23.0; IBM). Statistical analyses focused on the average response within a predefined midcentral region of interest (ROI; see inset in Fig. 2*A*), which is typical of auditory responses in EEG (Picton et al., 1974). The ROI included 92 electrodes that lay within a 6 cm radius from Cz. Fisher's z -transformation was applied to individual ITPC values before statistical analysis. Hypothesis testing consisted of independent comparisons between each of the speech conditions (intelligible, scrambled, and foreign language) and the sham (no stimulation) condition. Hypotheses were tested via paired t tests for all comparisons after verifying normality via Kolmogorov–Smirnov tests. Otherwise, a Wilcoxon rank-sum test was used. To ensure that effects in wakefulness were exclusive for the a priori frequencies of interest (0.5, 1, 2, and 4 Hz), we tested the initial ITPC results statistically on a wider range of 12 frequencies (every 0.5 Hz between 0.5 and 6 Hz). To account for multiple comparisons (3 speech conditions vs sham \times 12 frequencies = 36 comparisons), we controlled the false discovery rate (FDR) using a q value of 0.05 (Benjamini and Yekutieli, 2001). After confirming that significant responses in wakefulness are only observed at the a priori frequencies of interest, subsequent comparisons of sleep and wakefulness were restricted to those frequencies and corrected for 12 comparisons (3 speech conditions vs sham \times 4 frequencies) via FDR correction at $q = 0.05$.

Given the differences in ongoing spontaneous activity across states (which was substantially reduced by the CSD transformation used here, but not entirely equated; see Fig. 3*D*), direct comparisons across arousal states were performed via two-way repeated-measures ANCOVA tests including the SNR at each state as a covariate. This analysis focused on the intelligible speech condition comparing the response at each frequency of interest (1, 2, and 4 Hz) across states (wakefulness vs REM/NREM). We also tested whether the response at the acoustic/syllabic level (4 Hz) differed across states (wakefulness vs sleep) and speech conditions (intelligible, scrambled, foreign speech). Note that these ANCOVAs were performed separately when comparing wakefulness versus REM/NREM given the lower number of participants with REM sleep.

Results

Neuronal speech tracking during wakefulness is evident in the entire acoustic–linguistic hierarchy during intelligible speech

To validate the utility of the CHT paradigm with scalp EEG, we first assessed the ITPC of cortical activity during wakefulness across a range of 12 frequencies between 0.5 and 6 Hz (0.5 Hz intervals) within a midcentral ROI. ITPC at the acoustic/syllabic rate of 4 Hz was significant in all speech conditions (all t tests vs sham condition $p < 10^{-4}$) and remained significant after FDR correction at $q = 0.05$ (see Table 1 for detailed p -values; Fig. 2*A*). ITPC at frequencies corresponding to higher linguistic structures (2, 1, and 0.5 Hz) was significant only for the intelligible speech condition (word level: $p < 0.001$; phrase level: $p = 0.007$; sentence level: $p = 0.001$; remained significant after FDR correction; Fig. 2*A*). None of the unintelligible speech conditions elicited significant ITPC at any of these frequencies (Table 1). Importantly, significant ITPC was only observed at the frequencies of

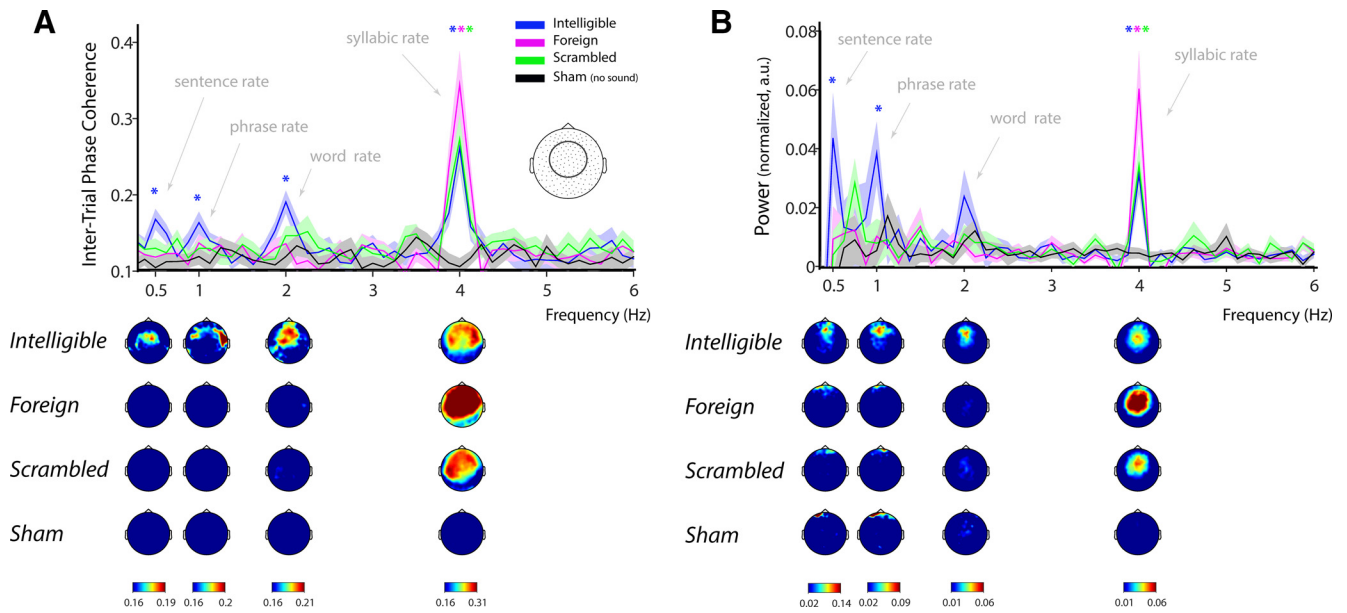


Figure 2. Neural tracking of hierarchical linguistic structures during wakefulness. **A**, Top, ITPC of EEG activity at a central ROI (inset, top right) for all speech conditions (blue, intelligible; green, scrambled; pink, foreign; black, sham). Asterisks represent statistically significant differences ($p < 0.02$), which remained significant after FDR correction at $q = 0.05$, between each speech condition and the sham condition. Shaded highlights denote SEM across participants ($n = 21$). All speech conditions elicited greater ITPC at 4 Hz compared with sham, whereas only the intelligible condition was associated with significant ITPC at frequencies representing linguistic structures (words, phrases, and sentences; 2, 1, and 0.5 Hz, respectively). Bottom, Mean ITPC topographies for each condition (rows) and frequency of interest (column). **B**, Top, Evoked power spectrum of EEG activity at a central ROI (same as **A**) reveals increased power at the syllabic/acoustic rate (4 Hz) for all speech stimulations and significant power only for the intelligible speech at 0.5, 1, and 2 Hz, corresponding to high-level linguistic structures.

Table 1. Statistical significance of ITPC analyses across states

	Sentence level (0.5 Hz)	Phrase level (1 Hz)	Word level (2 Hz)	Syllable level (4 Hz)
Wakefulness				
Intelligible speech vs sham	<0.001	0.007	0.001	<0.001
Scrambled speech vs sham	0.125	0.302	0.044	<0.001
Foreign speech vs sham	0.360	0.252	0.351	<0.001
Wakefulness ($n = 17$, NREM subset)				
Intelligible speech vs sham	0.270	0.019	0.004	<0.001
Scrambled speech vs sham	0.945	0.972	0.158	<0.001
Foreign speech vs sham	0.371	0.783	0.654	<0.001
Wakefulness ($n = 13$, REM subset)				
Intelligible speech vs sham	0.462	0.010	0.002	<0.001
Scrambled speech vs sham	0.573	0.238	0.412	<0.001
Foreign speech vs sham	0.992	0.028	0.725	<0.001
NREM sleep				
Intelligible speech vs sham	0.967	0.487	0.106	<0.001
Scrambled speech vs sham	0.450	0.427	0.192	<0.001
Foreign speech vs sham	0.351	0.968	0.177	<0.001
REM sleep				
Intelligible speech vs sham	0.462	0.906	0.033	<0.001
Scrambled speech vs sham	0.795	0.603	0.138	0.001
Foreign speech vs sham	0.682	0.893	0.069	<0.001

Table includes p -values comparing each speech condition (intelligible/scrambled/foreign) with the sham condition (no stimulation) separately for each frequency of interest. Comparisons were performed via paired t test (whenever normality was confirmed) or Wilcoxon (otherwise). Statistically significant p -values after FDR correction at $q = 0.05$ are marked in bold font; nonsignificant p -values are unbolded. Rows (top to bottom) mark full analysis during wakefulness ($n = 21$), wakefulness dataset matched to NREM sleep ($n = 17$, fewer trials), wakefulness dataset matched to REM sleep ($n = 13$, fewer trials), NREM sleep, and REM sleep. Note that the response at the syllabic/acoustic level (4 Hz) was robustly significant across all conditions, whereas significant responses at frequencies corresponding to higher-order linguistic structures were only observed for intelligible speech during wakefulness.

interest corresponding to predetermined linguistic structures in the stimuli ($p > 0.06$ at all other frequencies).

We repeated the same analysis for the EEG evoked power within the same central ROI (Fig. 2B). Similar to ITPC, normalized EEG power at the acoustic/syllabic rate of 4 Hz was

significant in all speech conditions ($p < 0.002$) and remained significant after FDR correction, whereas significant power at frequencies representing linguistic parsing was observed only for intelligible speech for most linguistic levels (sentence level: $p = 0.004$; phrase level: $p = 0.012$) and remained significant after FDR correction (word level: $p = 0.21$, n.s.). Corresponding scalp topographies were consistent with those observed for ITPC (Fig. 2B, bottom).

Overall, our ITPC and power analyses during wakefulness converge with previous results (Ding et al., 2016), indicating that, for intelligible speech, neural tracking is evident throughout the linguistic hierarchy, whereas for unintelligible speech, neural responses can be attributed to the acoustic structure.

Sleep preserves auditory responses but disrupts high-order linguistic parsing

Continuous overnight recordings lasted 9 h and 34 min on average (± 75 min). We verified that normal sleep was preserved in the presence of speech stimulation. Figure 3 illustrates the sleep architecture and spectral content of EEG activity, demonstrating all the established hallmarks of the different vigilance states, including alpha (8–10 Hz) activity during quiet wakefulness, sleep spindle (10–15 Hz) and slow-wave (<4 Hz) activities during N2/N3 sleep, and diffuse theta (6–9 Hz) during REM sleep. Awakenings associated with speech stimulation were rare (and even if such events were present and not observed, these trials would be tagged as wake or N1 sleep, so any differences reported here would constitute a lower bound). Sleep parameters (Fig. 3E) at the group level were in accordance with typical values for healthy young adults (Carskadon and Dement, 2005). In the morning debriefing, all participants reported being well rested (slept very well, not tired). Most participants vaguely recalled hearing the stimuli a few times after going to sleep, but all confirmed that this did not interfere with sleep quality because they returned back to sleep immediately. Therefore, intermittent

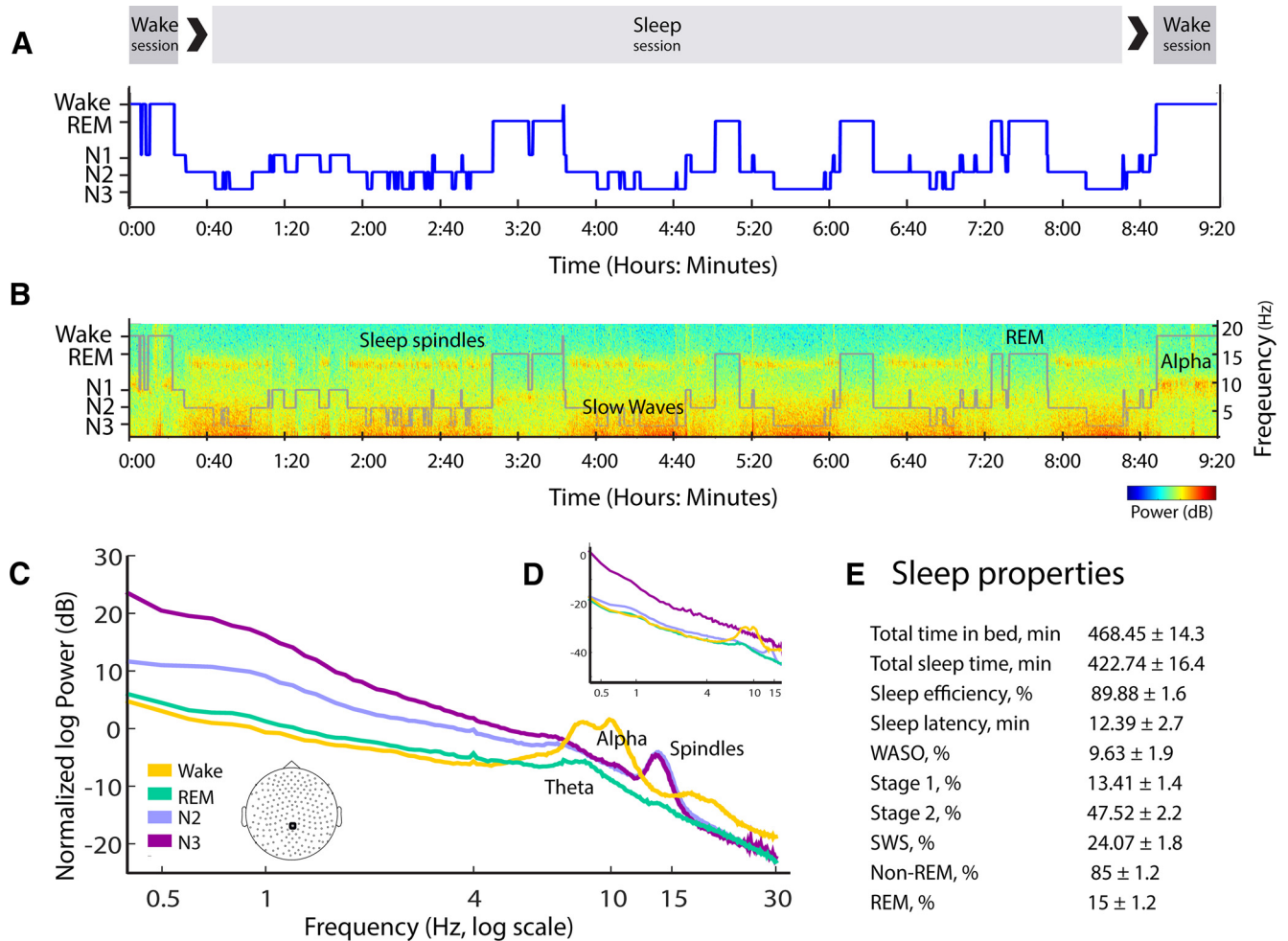


Figure 3. Sleep properties during overnight auditory stimulation. **A**, Representative hypnogram ($n = 1$) marking the time course of sleep/wake states (shown on left side of y-axis). **B**, Representative time–frequency representation of scalp EEG (Pz) during a full-night sleep study ($n = 1$). Warm colors mark increased power in specific time–frequency windows (frequency scale on the right). Superimposed hypnogram (gray trace, same as **A**). Note that wakefulness is associated with increased alpha power (8–12 Hz); N2 and N3 NREM sleep are characterized by increased slow waves (<4 Hz) and spindles (10–15 Hz); and REM sleep exhibits diffuse theta activity (6–9 Hz). **C**, Mean power spectrum of different wakefulness and sleep states ($n = 21$; yellow, wake; green, REM; light purple, N2; dark purple, N3). **D**, Mean power spectrum of different wakefulness and sleep states after performing CSD on the data. Note that the CSD procedure has substantially reduced the overall power at low frequencies for both REM and N2, but not for N3. **E**, Overnight measures of sleep ($n = 21$) indicate that sleep was largely normal for a first night in a sleep laboratory. All values are expressed as mean ± SEM. Percentage values are expressed per total sleep time excluding initial and final blocks of stimulation during wakefulness. Sleep efficiency corresponds to total sleep time per time in bed. Sleep latency is time to N2 sleep. WASO, Waking after sleep onset; SWS, slow-wave sleep.

speech stimulation did not exert significant effects on sleep architecture or subjective measures.

We proceeded to reanalyze EEG activity as a function of sleep and wakefulness states after applying a CSD spatial filter to reduce the effects of slow-wave activity during sleep (see Materials and Methods and Fig. 4). We focused our analysis on comparing wakefulness with N2 sleep ($n = 17$) and comparing wakefulness with REM sleep ($n = 13$), equating the number of trials and participants in each comparison to ensure identical statistical power. We did not analyze responses during N1 sleep because it was ambiguous and rare ($13.4 \pm 1.4\%$ of sleep time) or during N3 sleep (in which slow-wave activity precluded analysis; see Materials and Methods).

Our reanalysis of the wakefulness data (performed after CSD in the subset of participants with adequate sleep data) confirmed that, even in these reduced datasets, all speech conditions elicited a significant acoustic/syllabic-rate response at 4 Hz ($p < 10^{-3}$ for both datasets). Moreover, intelligible speech ITPC remained significant at the word and phrase levels in these reduced datasets (Fig. 4, top, Table 1). However, the sentential rate of 0.5 Hz was

no longer significant in either subset, probably due to the lower power of this analysis and/or the removal of low frequencies by the CSD.

In the sleep conditions, significant ITPC at the acoustic/syllabic rate of 4 Hz was observed in both sleep states for all stimuli (NREM: $p < 10^{-5}$; REM: $p < 10^{-3}$) and remained significant after FDR correction). Furthermore, we compared the 4 Hz (acoustic/syllabic) response for all audible conditions during wakefulness versus each sleep state directly using repeated-measures ANCOVA with SNR differences (sleep–wake) as a covariate. This comparison did not reveal significant differences in the acoustic/syllabic response across states (wakefulness vs REM: $F_{(2,10)} = 0.063$, $p = 0.940$, n.s.; wakefulness vs NREM: $F_{(2,14)} = 0.624$, $p = 0.550$, n.s.), suggesting no significant change in the acoustic response across states.

Importantly, despite robust activity at the acoustic level, none of the speech conditions during sleep elicited significant ITPC response versus sham at frequencies corresponding to high-level linguistic structures (Fig. 4A, B; Table 1). In addition, we compared the responses to intelligible speech during wakefulness ver-

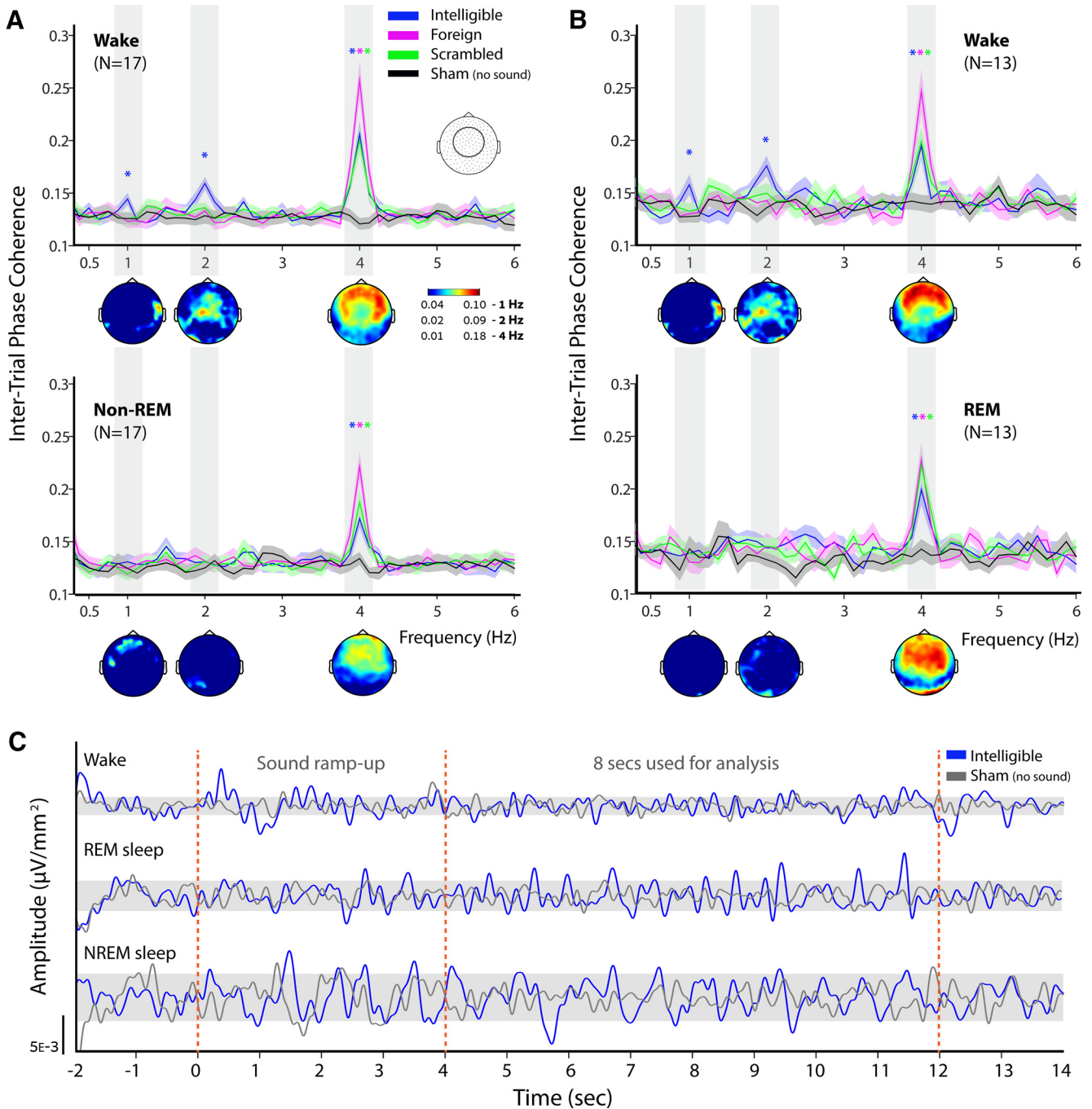


Figure 4. Sleep preserves auditory responses but disrupts high-order linguistic parsing. **A**, ITPC spectrum for wakefulness (top) and NREM sleep (bottom; $n = 17$) within a midcentral ROI (inset) for all speech conditions (blue, intelligible; green, scrambled; pink, foreign; black, sham). Asterisks mark statistically significant differences ($p < 0.02$), which remained significant after FDR correction at $q = 0.05$, between each speech condition and the sham condition. Shaded highlights denote SEM across participants. Below each frequency and state are corresponding normalized ITPC topographies (see Materials and Methods) in response to intelligible speech. Note that NREM sleep is associated with significant ITPC at the syllabic/acoustic rate (4 Hz) for all conditions, but disrupted ITPC at frequencies of high-level linguistic structures. **B**, ITPC spectrum for wakefulness (top) and REM sleep (bottom; $n = 13$). ROI, colors, gray shading, and topographies as in **A**. Note that REM sleep is associated with significant ITPC at the syllabic/acoustic rate (4 Hz) for all conditions, but no significant ITPC at frequencies of high-level linguistic structures. **C**, Grand-mean ERPs from the midcentral ROI in wakefulness (top), REM sleep (middle), and NREM sleep (bottom) during 12 s trials and ± 2 s pretrial and posttrial. Blue, intelligible speech; gray, sham condition. Gray shading represents the root-mean-square (RMS) of ongoing activity (without stimulation). Note that, in all states, the magnitude of evoked responses exceeds ongoing activity.

sus each sleep state at frequencies of interest (1, 2, and 4 Hz) using ANCOVA with SNR difference as a covariate. This analysis revealed significant state \times frequency interaction effects (wakefulness vs REM: $F_{(2,10)} = 5.21, p = 0.028$; wakefulness vs NREM: $F_{(2,14)} = 4.35, p = 0.034$), consistent with the comparisons of each speech condition versus the sham condition performed within each state (Table 1).

Altogether, our results suggest that, whereas both NREM and REM sleep preserve acoustic responses, neural tracking of higher-order linguistic levels within a speech stream is not evident.

Discussion

The main novel result reported here is that, during sleep, basic neural encoding of acoustic features of speech persists, whereas

parsing of higher-order linguistic structures, evident for intelligible speech during wakefulness, is disrupted. These findings further our understanding of the capacity and limits of cortical processing during sleep and demonstrate the effectiveness of the CHT approach for probing high-level cognitive processes covertly in unresponsive states.

Studying linguistic processing during wakefulness using CHT

Our results in wakefulness are consistent with Ding et al. (2015), showing that cortical activity concurrently tracks the time course of linguistic structures at multiple hierarchical levels when speech is intelligible. For unintelligible speech, cortical responses are only evident at the acoustic/syllabic rate. Our results extend the original findings in several ways. First, using bisyllabic words allowed distinguishing between the syllabic (4 Hz) and word rate (2 Hz), affording further separation between the acoustic and linguistic aspects of speech. Second, our design did not include an explicit task, demonstrating that passive listening is sufficient for revealing hierarchical neural tracking of linguistic structures. Third, our results demonstrate that one can effectively quantify hierarchical linguistic parsing with noninvasive, inexpensive, portable, and readily available EEG. These extensions pave the way for using the CHT approach in a wide array of settings, ranging from language acquisition to studies of residual cognitive processing in various clinical populations.

Depth of speech processing during sleep

The current findings make a substantial contribution to our understanding of the extent of language processing during sleep. We found that the neural response at the syllabic rate, which likely reflects basic auditory representation of the acoustic envelope, was comparable across wakefulness, REM, and NREM sleep. This is consistent with previous studies supporting preserved responses in low-level auditory cortex during sleep (Issa and Wang, 2008; Nir et al., 2015). At the same time, parsing of higher-order linguistic structures is disrupted during sleep, implying the existence of a functional “bottleneck” precluding full processing of continuous speech.

Before discussing broader implications, we address an important methodological caveat: Is the lack of observable low-frequency peaks during sleep due to increased background noise leading to poor sensitivity rather than a lack of high-level speech parsing? This potential criticism applies mainly to NREM sleep because spontaneous activity during REM sleep is wake-like. Although applying CSD transformation effectively reduced slow widespread activity in NREM sleep (Fig. 3D), it was nevertheless still higher than in wakefulness. However, we believe that this does not account for our results for two reasons. First, our experimental design deliberately focused on testing different speech conditions within each state, in which SNR and other signal properties are matched. Second, we conducted several analyses taking into account SNR differences across states. These showed that, although the 4 Hz acoustic responses did not differ significantly across states, responses at the word and sentence level were found for intelligible speech only during wakefulness, as indicated by significant interactions between state \times frequency. Nevertheless, we acknowledge that, given the dominance of low-frequency background activity during NREM sleep (as in other nonresponsive states such as anesthesia and vegetative states), results should be interpreted with great caution and appropriate controls applied in future studies.

What is the nature of the functional bottleneck observed during sleep between basic auditory processing and high-level lin-

guistic parsing? Previous ECoG results using the CHT paradigm attribute phrasal and sentential responses to nonsensory areas involved in language processing; for example, the left IFG and bilateral TPJ regions. In contrast, syllabic rate responses were found primarily in auditory cortex (Ding et al., 2016). The lack of evidence for neural tracking of high-level speech structures during sleep suggests that it disrupts efficient signal propagation from auditory cortex to higher cortical regions. Along these lines, intercortical connectivity during sleep is restricted upon brief electromagnetic perturbation (Massimini et al., 2005, 2007; Pigorini et al., 2015). Similarly, several fMRI studies demonstrated robust attenuation of speech responses in IFG and frontal regions during sleep (Portas et al., 2000; Dehaene-Lambertz et al., 2002; Wilf et al., 2016). Together with the current findings, these studies imply that speech processing during sleep is limited to low-level acoustic processing, with comparable neural responses for intelligible and unintelligible speech.

However, other studies report evidence of residual semantic processing during sleep upon presentation of single words (Kouider et al., 2014; Andrillon et al., 2016), word pairs (Brualla et al., 1998), and short sentences (Ibáñez et al., 2006; Daltrozzo et al., 2012), although such processing is generally weaker and with altered time dynamics compared with wakefulness (Brualla et al., 1998; Perrin et al., 2002). In an attempt to reconcile these seemingly contradictory results, we suggest that the functional bottleneck for speech processing during sleep is not semantic analysis *per se*, but rather the mediating process of segmentation. One key difference between this and previous studies is the use of continuous speech rather than single words or short sentences. This not only allowed us to overcome serious methodological limitations such as the pervasiveness of K-complexes, but also to tap into the process of word segmentation, which constitutes a prerequisite for continuous speech comprehension. Indeed, speech processing relies critically on accurately parsing the ongoing acoustic stream input into discrete and meaningful linguistic units (Hickok et al., 1993; Mattys, 1997; Giraud and Poeppel, 2012; Doelling et al., 2014). Parsing itself is a hierarchical highly demanding cognitive process (Greenberg et al., 2004; Ghitza, 2012) that involves matching bottom-up cues with existing lexical, syntactic, and semantic representations (Traxler, 2014), as well as predictive and anticipatory processes (Arnal et al., 2011; DeLong et al., 2014). The current findings suggest that, at minimum, it is the process of word segmentation that is compromised during sleep, a process that requires ongoing analysis of the continuous global stream of input (Strauss et al., 2015; Tononi et al., 2016). It thus remains possible that, when phrases or sentences are presented in a discrete format (as in most previous studies), individual words can be identified and processed fully up to a semantic level even during sleep.

Another critical aspect of continuous speech parsing is integrating information across multiple time scales (Rosen, 1992; Greenberg et al., 2003; Ghitza, 2012; Zion Golumbic et al., 2013), allowing short-scale information (phonemes/syllables) to be combined to create lexical units (e.g., words) and higher-order structures (phrases and sentences). Such integration requires short-term memory buffers to sustain information for longer durations; for example, to handle larger “temporal receptive windows” (Lerner et al., 2011; Luo and Poeppel, 2012; Chait et al., 2015). Notably, previous studies have shown that neural activity during sleep is restricted to short timescales in the range of hundreds of milliseconds, preventing the emergence of long-lasting causal interactions (Pigorini et al., 2015; Strauss et al., 2015). The proposed restriction on brain activity during sleep to

short time intervals may also contribute to the diminished hierarchical parsing of continuous speech observed here.

One important limitation of the current study is that prosodic cues were purposefully removed from the speech material, allowing us to probe grammatical speech parsing regardless of correlated acoustic variations. However, because syntactic parsing of natural speech also benefits from prosody (Eckstein and Friederici, 2006), it remains to be tested whether prosody can facilitate continuous speech parsing during sleep.

Sleep and inattention

Analogous results to those found here during sleep have been reported recently for unattended speech, for which neural tracking is robust in auditory cortex, but substantially reduced in higher-order language areas (Zion Golumbic et al., 2013; Rimele et al., 2015). Along this line, acoustic features of unattended speech affect behavior more than high-level (semantic) features (Ellermeier et al., 2015; Wöstmann and Obleser, 2016). Additional similarities between sleep and inattention exist at the behavioral level, where unattended speech typically cannot be overtly recalled (Lachter et al., 2004), although some personally relevant words occasionally capture attention (Cherry, 1953; Moray, 1958, 1959; Bentin et al., 1995; Wood and Cowan, 1995), as is also found during sleep. These similarities between speech processing during sleep and inattention may carry broader implications as to the nature of functional “bottlenecks” in speech processing (Broadbent, 1958; Treisman, 1969), constituting an intriguing topic for future research.

Conclusions

To the best of our knowledge, this study is the first to investigate hierarchical parsing of continuous speech during sleep. We found that bottom-up auditory processing is preserved in sleep and comparable to that found in wakefulness. In sharp contrast, neural tracking of high-order linguistic structures—words, phrases, and sentences—is disrupted in sleep in a manner similar to unattended or unintelligible speech during wakefulness. Current results suggest that parsing of continuous speech, which requires integration across multiple time scales, matching of bottom-up input to stored linguistic representations, and top-down predictive coding, may not be possible without overt attention and consciousness. Our results imply a functional barrier between auditory sensation and linguistic processing, a barrier that may be essential to ensure preservation of sleep in the face of external events and to support its functions. This study sets the ground toward studying residual speech processing across states of consciousness, anesthesia, neurodegeneration, development, and language disorders.

References

Andrillon T, Nir Y, Cirelli C, Tononi G, Fried I (2015) Single-neuron activity and eye movements during human REM sleep and awake vision. *Nat Commun* 6:7884. [CrossRef Medline](#)

Andrillon T, Poulsen AT, Hansen LK, Léger D, Kouider S (2016) Neural markers of responsiveness to the environment in human sleep. *J Neurosci* 36:6583–6596. [CrossRef Medline](#)

Arnal LH, Wyart V, Giraud AL (2011) Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nat Neurosci* 14:797–801. [CrossRef Medline](#)

Bastuji H, Perrin F, Garcia-Larrea L (2002) Semantic analysis of auditory input during sleep: studies with event related potentials. *Int J Psychophysiol* 46:243–255. [CrossRef Medline](#)

Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 29:1165–1188. [CrossRef](#)

Bentin S, Kutas M, Hillyard SA (1995) Semantic processing and memory for

attended and unattended words in dichotic listening: behavioral and electrophysiological evidence. *J Exp Psychol Hum Percept Perform* 21:54–67. [CrossRef Medline](#)

Berens P (2009) CircStat: A MATLAB toolbox for circular statistics. *Journal of Statistical Software* 31:1–21.

Blume C, Del Giudice R, Lechinger J, Wislowska M, Heib DP, Hoedlmoser K, Schabus M (2017) Preferential processing of emotionally and self-relevant stimuli persists in unconscious N2 sleep. *Brain Lang* 167:72–82. [CrossRef Medline](#)

Boersma P, Weenink D (2016) Praat: doing phonetics by computer [Computer program]. Version 5.4.04, retrieved 30 December 2014.

Broadbent DE (1958) Selective listening to speech. In: *Perception and Communication*, pp. 11–35. Oxford: Pergamon Press.

Brualla J, Romero MF, Serrano M, Valdizán JR (1998) Auditory event-related potentials to semantic priming during sleep. *Electroencephalogr Clin Neurophysiol* 108:283–290. [CrossRef Medline](#)

Carskadon MA, Dement WC (2005) Normal human sleep: an overview. In: *Principles and practice of sleep medicine* (Kryger M, Roth T, Dement W, eds.), pp 16–26. St. Louis: Elsevier Saunders.

Chait M, Greenberg S, Arai T, Simon JZ, Poeppel D (2015) Multi-time resolution analysis of speech: evidence from psychophysics. *Front Neurosci* 9:214. [CrossRef Medline](#)

Cherry EC (1953) Some experiments on the recognition of speech, with one and with 2 ears. *J Acoust Soc Am* 25:975–979. [CrossRef](#)

Colrain IM (2005) The K-complex: a 7-decade history. *Sleep* 28:255–273. [CrossRef Medline](#)

Colrain IM, Campbell KB (2007) The use of evoked potentials in sleep research. *Sleep Med Rev* 11:277–293. [CrossRef Medline](#)

Daltrozzo J, Claude L, Tillmann B, Bastuji H, Perrin F (2012) Working Memory Is Partially Preserved during Sleep. *PLoS One* 7.

Dehaene-Lambertz G, Dehaene S, Hertz-Pannier L (2002) Functional neuroimaging of speech perception in infants. *Science* 298:2013–2015. [CrossRef Medline](#)

DeLong KA, Troyer M, Kutas M (2014) Pre-processing in sentence comprehension: Sensitivity to likely upcoming meaning and structure. *Lang Linguist Compass* 8:631–645. [CrossRef Medline](#)

Ding N, Melloni L, Zhang H, Tian X, Poeppel D (2016) Cortical tracking of hierarchical linguistic structures in connected speech. *Nat Neurosci* 19:158–164. [CrossRef Medline](#)

Doelling KB, Arnal LH, Ghitza O, Poeppel D (2014) Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *Neuroimage* 85:761–768. [CrossRef Medline](#)

Eckstein K, Friederici AD (2006) It's early: event-related potential evidence for initial interaction of syntax and prosody in speech comprehension. *J Cogn Neurosci* 18:1696–1711. [CrossRef Medline](#)

Ellermeier W, Kattner F, Ueda K, Doumoto K, Nakajima Y (2015) Memory disruption by irrelevant noise-vocoded speech: effects of native language and the number of frequency bands. *J Acoust Soc Am* 138:1561–1569. [CrossRef Medline](#)

Friederici AD, Singer W (2015) Grounding language processing on basic neurophysiological principles. *Trends Cogn Sci* 19:329–338. [CrossRef Medline](#)

Ghitza O (2012) On the role of theta-driven syllabic parsing in decoding speech: intelligibility of speech with a manipulated modulation spectrum. *Front Psychol* 3:238. [CrossRef Medline](#)

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

Greenberg S, Carvey H, Hitchcock L, Chang S (2003) Temporal properties of spontaneous speech—a syllable-centric perspective. *J Phon* 31:465–485. [CrossRef](#)

Greenberg S, Ainsworth W a, Popper AN, Fay RR (2004) *Speech processing in the auditory system*. New York: Springer.

Halász P (2016) The K-complex as a special reactive sleep slow wave: a theoretical update. *Sleep Med Rev* 29:34–40. [CrossRef Medline](#)

Hickok G, Zurif E, Canseco-Gonzalez E (1993) Structural description of agrammatic comprehension. *Brain Lang* 45:371–395. [CrossRef Medline](#)

Ibáñez A, López V, Cornejo C (2006) ERPs and contextual semantic discrimination: Degrees of congruence in wakefulness and sleep. *Brain Lang* 98:264–275. [CrossRef Medline](#)

Iber C, Ancoli-Israel S, Chesson A, Quan S (2007) *The AASM manual for the scoring of sleep and associated events: rules, terminology and techni-*

- cal specifications, Ed 1. Westchester, IL: American Academy of Sleep Medicine.
- Issa EB, Wang X (2008) Sensory responses during sleep in primate primary and secondary auditory cortex. *J Neurosci* 28:14467–14480. [CrossRef Medline](#)
- Kayser J, Tenke CE (2015) On the benefits of using surface Laplacian (current source density) methodology in electrophysiology. *Int J Psychophysiol* 97:171–173. [CrossRef Medline](#)
- Kouider S, Andrillon T, Barbosa LS, Goupil L, Bekinschtein TA (2014) Inducing task-relevant responses to speech in the sleeping brain. *Curr Biol* 24:2208–2214. [CrossRef Medline](#)
- Lachter J, Forster KI, Ruthruff E (2004) Forty-five years after Broadbent (1958): still no identification without attention. *Psychol Rev* 111:880–913. [CrossRef Medline](#)
- Langford GW, Meddis R, Pearson AJ (1974) Awakening latency from sleep for meaningful and non-meaningful stimuli. *Psychophysiology* 11:1–5. [CrossRef Medline](#)
- Lerner Y, Honey CJ, Silbert LJ, Hasson U (2011) Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J Neurosci* 31:2906–2915. [CrossRef Medline](#)
- Luo H, Poeppel D (2012) Cortical oscillations in auditory perception and speech: evidence for two temporal windows in human auditory cortex. *Front Psychol* 3:170. [CrossRef Medline](#)
- Massimini M, Ferrarelli F, Huber R, Esser SK, Singh H, Tononi G (2005) Breakdown of cortical effective connectivity during sleep. *Science* 309:2228–2232. [CrossRef Medline](#)
- Massimini M, Ferrarelli F, Esser SK, Riedner BA, Huber R, Murphy M, Peterson MJ, Tononi G (2007) Triggering sleep slow waves by transcranial magnetic stimulation. *Proc Natl Acad Sci U S A* 104:8496–8501. [CrossRef Medline](#)
- Mattys SL (1997) The use of time during lexical processing and segmentation: a review. *Psychon Bull Rev* 4:310–329. [CrossRef](#)
- McDonald DG, Schicht WW, Frazier RE, Shallenberger HD, Edwards DJ (1975) Studies of information processing in sleep. *Psychophysiology* 12:624–629. [CrossRef Medline](#)
- Moray N (1958) The effect of the relative intensities of dichotic messages in speech shadowing. *Lang Speech* 1:110–113.
- Moray N (1959) Attention in dichotic listening: affective cues and the influence of instructions. *Q J Exp Psychol* 11:56–60. [CrossRef](#)
- Nir Y, Tononi G (2010) Dreaming and the brain: from phenomenology to neurophysiology. *Trends Cogn Sci* 14:88–100. [CrossRef Medline](#)
- Nir Y, Vyazovskiy VV, Cirelli C, Banks MI, Tononi G (2015) Auditory responses and stimulus-specific adaptation in rat auditory cortex are preserved across NREM and REM sleep. *Cereb Cortex* 25:1362–1378. [CrossRef Medline](#)
- Nozaradan S, Peretz I, Missal M, Mouraux A (2011) Tagging the neuronal entrainment to beat and meter. *J Neurosci* 31:10234–10240. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Oswald I, Taylor AM, Treisman M (1960) Discriminative responses to stimulation during human sleep. *Brain* 83:440–453. [CrossRef Medline](#)
- Peelle JE, Johnsrude IS, Davis MH (2010) Hierarchical processing for speech in human auditory cortex and beyond. *Front Hum Neurosci* 4:51. [CrossRef Medline](#)
- Peña JL, Pérez-Perera L, Bouvier M, Velluti RA (1999) Sleep and wakefulness modulation of the neuronal firing in the auditory cortex of the guinea pig. *Brain Res* 816:463–470. [CrossRef Medline](#)
- Perrin F, Bastuji H, Garcia-Larrea L (2002) Detection of verbal discordances during sleep. *Neuroreport* 13:1345–1349. [CrossRef Medline](#)
- Picton TW, Hillyard SA, Krausz HI, Galambos R (1974) Human auditory evoked potentials. I: Evaluation of components. *Electroencephalogr Clin Neurophysiol* 36:179–190. [CrossRef Medline](#)
- Pigorini A, Sarasso S, Proserpio P, Szymanski C, Arnulfo G, Casarotto S, Fecchio M, Rosanova M, Mariotti M, Lo Russo G, Palva JM, Nobili L, Massimini M (2015) Bistability breaks-off deterministic responses to intracortical stimulation during non-REM sleep. *Neuroimage* 112:105–113. [CrossRef Medline](#)
- Portas CM, Krakow K, Allen P, Josephs O, Armony JL, Frith CD (2000) Auditory processing across the sleep-wake cycle: Simultaneous EEG and fMRI monitoring in humans. *Neuron* 28:991–999. [CrossRef Medline](#)
- Pratt H, Berlad I, Lavie P (1999) “Oddball” event-related potentials and information processing during REM and non-REM sleep. *Clin Neurophysiol* 110:53–61. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Rosen S (1992) Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos Trans R Soc Lond B Biol Sci* 336:367–373. [CrossRef Medline](#)
- Strauss M, Sitt JD, King JR, Elbaz M, Azizi L, Buiatti M, Naccache L, van Wassenhove V, Dehaene S (2015) Disruption of hierarchical predictive coding during sleep. *Proc Natl Acad Sci U S A* 112:E1353–E1362. [CrossRef Medline](#)
- Tononi G, Boly M, Massimini M, Koch C (2016) Integrated information theory: from consciousness to its physical substrate. *Nat Rev Neurosci* 17:450–461. [CrossRef Medline](#)
- Traxler MJ (2014) Trends in syntactic parsing: Anticipation, Bayesian estimation, and good-enough parsing. *Trends Cogn Sci* 18:605–611. [CrossRef Medline](#)
- Treisman AM (1969) Strategies and models of selective attention. *Psychol Rev* 76:282–299. [CrossRef Medline](#)
- Wilf M, Ramot M, Furman-Haran E, Arzi A, Levkovitz Y, Malach R (2016) Diminished auditory responses during NREM sleep correlate with the hierarchy of language processing. *PLoS One* 11:e0157143. [CrossRef Medline](#)
- Wood N, Cowan N (1995) The cocktail party phenomenon revisited: how frequent are attention shifts to one’s name in an irrelevant auditory channel? *J Exp Psychol Learn Mem Cogn* 21:255–260. [CrossRef Medline](#)
- Wöstmann M, Obleser J (2016) Acoustic detail but not predictability of task-irrelevant speech disrupts working memory. *Front Hum Neurosci* 10:538. [Medline](#)
- Wylie KP, Regner MF (2014) Large-scale network involvement in language processing. *J Neurosci* 34:15505–15507. [CrossRef Medline](#)
- Zion Golumbic EM, 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. [CrossRef](#)