The Brain Selectively Tunes to Unfamiliar Voices during Sleep

Mohamed S. Ameen, Dominik P.J. Heib, Christine Blume, and Manuel Schabus

The brain continues to respond selectively to environmental stimuli during sleep. However, the functional role of such responses, and whether they reflect information processing or rather sensory inhibition, is not fully understood. Here, we present 17 human sleepers (14 females) with their own name and two unfamiliar first names, spoken by either a familiar voice (FV) or an unfamiliar voice (UFV), while recording polysomnography during a full night of sleep. We detect K-complexes, sleep spindles, and microarousals, and assess event-related and frequency responses as well as intertrial phase synchronisation to the different stimuli presented during nonrapid eye movement (NREM) sleep. We show that UFVs evoke more K-complexes and microarousals than FVs. When both stimuli evoke a K-complex, we observe larger evoked potentials, more precise time-locking of brain responses in the delta band (1–4 Hz), and stronger activity in the high frequency (>16 Hz) range, in response to UFVs relative to FVs. Crucially, these differences in brain responses disappear completely when no K-complexes are evoked by the auditory stimuli. Our findings highlight discrepancies in brain responses to auditory stimuli based on their relevance to the sleeper and propose a key role for K-complexes in the modulation of sensory processing during sleep. We argue that such content-specific, dynamic reactivity to external sensory information enables the brain to enter a sentinel processing mode in which it engages in the important internal processes that are ongoing during sleep while still maintaining the ability to process vital external sensory information.

Key words: auditory stimulation; EEG; information processing; K-complexes; microarousals; sleep

Significance Statement

Previous research has shown that sensory processing continues during sleep. Here, we studied the capacity of the sleeping brain to extract and process relevant sensory information. We presented sleepers with their own names and unfamiliar names spoken by either an FV or a UFV. During NREM sleep, UFVs elicited more K-complexes and microarousals than FVs. By contrasting stimuli that evoked K-complexes, we demonstrate that UFVs evoked larger, more synchronized brain responses as well as stronger power at high frequencies (>16 Hz) relative to FVs. These differences in brain responses disappeared when no K-complexes were evoked. Our results suggest a pivotal role for K-complexes in the selective processing of relevant information during NREM sleep.

Introduction

During sleep, the brain continues to respond to auditory stimuli in a selective fashion (Portas et al., 2000; Andrillon et al., 2016; Blume et al., 2017, 2018). Previous studies have demonstrated, for instance, that the subject’s own name (SON) evokes stronger brain responses than other names during sleep (Oswald et al., 1960; Perrin et al., 1999; Pratt et al., 1999). Blume et al. (2018) showed that during all stages of sleep, brain responses to SON and other unfamiliar names (UNs) did not differ; however, names uttered by an unfamiliar voice (UFV) evoked stronger brain responses compared with a familiar voice (FV).

The discrepancy in brain responses to different stimuli implies the presence of an initial, presumably low-level, sensory processing during sleep that enables the brain to differentiate
between sensory signals (Blume et al., 2018). However, knowledge about the functions of such responses is still lacking. That is, the selective brain responses to specific sounds during sleep might reflect inhibitory processes that protect sleep from disruptions. Conversely, they might indicate higher, higher level processes that ensure the connectedness of the sleeping brain to the surrounding.

In this study, we investigated the purpose of such selective brain responses to sounds presented during nonrapid eye movement (NREM) sleep. We focused on sleep-specific events that have been previously linked to information processing, sensory inhibition, or both. That is, we focused on three cardinal sleep-specific electroencephalography (EEG) events, namely, the K-complex (KC), sleep spindles, and microarousals.

KCJs are ~1 Hz oscillations and a hallmark of Stage 2 NREM (N2) sleep (Loomis et al., 1938; Colrain, 2005; Halász, 2005). KCJs occur either spontaneously or in response to sensory stimuli. Spontaneous KCJs appear in the EEG signal as a well-defined sharp negative wave followed by a positive component with a total duration of at least 0.5 s (Rechtschaffen and Kales, 1968; Horí et al., 2001). Following sensory perturbation, KCJs appear to have two main components, a sharp negative (N) deflection at ~550 ms (N550) followed by a longer-lasting positive (P) wave at ~900 ms (P900; Bastien and Campbell, 1992; Cote et al., 1999; Colrain, 2005; Halász, 2005). Some studies have considered an early positive peak that appears ~200 ms (P200; Laurino et al., 2014, 2019) and another negative peak ~350 ms (N350; Bastien and Campbell, 1992; Cote et al., 1999) to be parts of the KC, albeit these components can occur without a KC being elicited. Relevant stimuli have a higher propensity to trigger KCJs (Halász, 2005). Theories suggest that KCs can serve both sleep-protecting as well as arousal-inducing processes (Halász, 2005; Jahnke et al., 2012; Forget et al., 2011; Laurino et al., 2014; Blume et al., 2017, 2018; Legendre et al., 2019; Latrelle et al., 2020).

Sleep spindles are also characteristic of N2 sleep. Spindles are thalamocortical oscillations of 11–15 Hz that last ~0.5–2 s (De Gennaro and Ferrara, 2003; Fernandez and Lüthi, 2020) and can be triggered by sensory stimuli (Antony and Paller, 2017). They have been repeatedly shown to inhibit sensory processing during sleep (McCormick and Bal, 1994; Schabus et al., 2012; Blume et al., 2018; Fernandez and Lüthi, 2020). However, some work challenges this notion (Sela et al., 2016) and even associates spindles with the processing of memory-related sounds presented during NREM sleep (Cairney et al., 2018).

Finally, microarousals are abrupt shifts in the EEG signal toward theta, alpha, and/or high beta (>$16$ Hz) frequencies (Halász et al., 1979; American Sleep Disorders Association, 1992; Halász et al., 2004) that appear in all sleep stages and are considered windows of information processing during sleep (Halász et al., 2004; Halász, 2005; dos Santos Lima et al., 2019). Microarousals are usually preceded by KCs (Colrain, 2005; Halász, 2005), yet they have been shown to be correlated with a lower incidence of sleep spindles in the preceding 10 s of EEG signal (Ehrhart et al., 1981).

Here, we reanalyzed the dataset used in Blume et al. (2018), who recorded polysomnography while presenting SONs and two UNs spoken by either an FV or a UFV during a whole night of sleep. We detected KCJs, spindles, and microarousals in response to these sounds during NREM sleep and hypothesized that the selective auditory-evoked responses support the extraction and processing of relevant sensory information.

### Materials and Methods

#### Participants

We recruited 20 healthy participants with no reported history of neurological or psychological problems as well as no reported sleep disorders. However, one participant dropped out after the adaptation night, and we had to exclude two participants because of technical problems during EEG acquisition. Therefore, we performed the analyses we report here on 17 participants (14 females) with a median age of $22.6 \pm 2.3$ years. Before beginning the experiment, all participants signed written informed consents. The experiment was approved by the Ethics Committee of the University of Salzburg.

#### Experimental design

Before the start of the experiment, participants were advised to maintain a regular sleep/wake cycle (~8 h of sleep) for at least 4 d, which we monitored via actigraphy (Fig. 1A). Subsequently, participants spent two nights in the sleep laboratory of the University of Salzburg. The first night was an adaptation night, during which we recorded polysomnography (PSG) data with no auditory stimulation. The second night was an experimental night, during which we recorded PSG data while presenting sounds via loudspeakers throughout the night. In both nights, participants were tested during wakefulness before and after sleep. Briefly, the wakefulness testing consisted of two sessions, a passive-listening session and an active-listening session. Passive listening entails the participants’ listening to the repeatedly presented auditory stimuli, whereas active listening means they had to count the number of presentations of one specific stimulus chosen by the experimenters. Before the wakefulness testing, participants were stimulated with either a bright (blue enriched) light or an inactive (sham) light for 1 h. The order of the light-stimulation conditions was counterbalanced between the adaptation and the experimental nights across participants. However, the light condition is irrelevant to this study as we grouped our data over both light conditions. For the purpose of this article, we focus primarily on the sleep part of the experimental night. For more details on the wakefulness part of the experiment, please refer to Blume et al. (2018). During the experimental night, participants went to bed around their habitual bedtime (8:30 P.M. to 11:30 P.M.). Time in bed (TIB) was ~8 h. After 8 h TIB, we waited for light NREM or REM sleep before waking up the participants (median sleep duration = 480 ± 2.5 min). The auditory stimulation started directly after participants went to bed and continued throughout the whole night. We presented auditory stimuli continuously for 90 min (Stimulation periods) then paused the presentation for 30 min (No-stimulation periods) to allow for periods of undisturbed sleep. This resulted in a 120 min cycle that we repeated four times throughout the night (Fig. 1B).

#### Stimuli

We presented six different auditory stimuli (Fig. 1C) that we personalized for each participant. The stimuli were the SON and two UNs spoken by either an FV or a UFV. An FV was the voice of someone close to the participant, for example, one of the parents. An UFV was the voice of someone unknown to the participant. We did not control for the sex of the voices, but they were matched; that is, the familiar and unfamiliar voices were always both either male or female. We chose UNs that matched the SON in the number of syllables and the frequency of occurrence in the population. The volume for stimulus presentation was adjusted individually for each participant so that the participant could clearly hear the stimulus and still be able to fall asleep. Each stimulus was presented 690 times, and the mean duration was 808 ± 110 ms. Stimuli were presented in a pseudorandom order, and no stimulus was presented twice in a row. The interstimulus intervals during sleep were jittered between 2800 and 7800 ms in 500 ms steps. Stimulus preprocessing, that is, denoising and normalization, was performed using Audacity software (https://audacityteam.org/). Stimulus delivery was controlled by MATHLAB (MathWorks).

#### Brain data acquisition

We recorded ongoing brain activity using a high-density EEG 256-channel GSN HydroCel Geodesic Sensor Net (Electrical Geodesics) and a Net...
Amps 400 amplifier. The PSG recordings included two electrooculography (EOG) and two chin electromyography (EMG) channels. Data were acquired at a sampling rate of 250 Hz, and Cz served as the online reference.

Sleep staging

Sleep staging was performed on 30 s epochs using the computer-assisted sleep classification system developed by the SIESTA Group (Somnolyzer 24 × 7; Anderer et al., 2005, 2010) following the standard criteria recommended by the American Association for Sleep Medicine (Iber, 2007). We have previously shown that the level of agreement between this algorithm and expert human scorers is similar to the level of agreement between human experts (Ameen et al., 2019).

The detection of sleep microstructures

KC detection. We detected KCs automatically with a wavelet-detection algorithm that was developed by the SIESTA Group. The development and validation procedures have been described in detail in Parapatics et al. (2015) and Schwarz et al. (2017), respectively. Briefly, 12 experienced human scorers visually scored KCs in 873 epochs of 10 min of PSG recordings from 189 control subjects and 90 patients. The features of the visually scored KCs were used to set the criteria for detection as well as to create a template KC that worked as a gold standard for the automatic detection. The detection itself is a two-step process; first, the algorithm detects possible KCs via an approach that combines a matched-filtering detection method and a slow-wave detection method (Woertz et al., 2004). Accordingly, the detection criteria for possible KCs were the following: (1) a minimum negative-to-positive peak-to-peak amplitude of 50 µV and (2) a duration between 480 and 1500 ms.

Second, all possible KCs are matched to the prototypical KC template via wavelet analysis, and the results are submitted to a linear discriminant analysis (LDA) to select only real KCs. For our analysis, we considered real KCs to be events that have an LDA score (how likely a specific EEG segment is a KC) of 0.8 or higher. This LDA score corresponds to 61.87 ± 6.9% of all detected KCs and a mean correlation to the template of 0.87 ± 0.007 over all subjects and is a compromise between reliable detections and a sufficient number of detected events for our analyses. Note that an LDA score of 1.7 reflects 98% detection specificity. Before running the detection algorithm, raw data were downsampled to 128 Hz and rereferenced to the contralateral mastoid. We detected KCs at C3 and C4. We only report results from C3 as the detections were similar between C3 and C4. We only considered events that occurred during N2 and N3 (Slow-wave) sleep and fulfilled the standard criteria for KC detection (Rechtschaffen and Kales, 1968; Hori et al., 2001). For N3 detections, however, we applied a stricter amplitude criterion as we only selected events with a peak-to-peak amplitude of 75 µV or higher (Cote et al., 1999; Nir et al., 2011). We marked the start of KC events as...
the point of the negative-going zero crossing of the signal before the negative peak. We defined evoked KCs as those events that occurred (started) in the 2000 ms poststimulus-onset window.

Figure 2 demonstrates the LDA distribution of the detected KCs and contains some examples of the detected events.

**Spindle detection.** Sleep spindles were detected using an algorithm developed by the SIESTA Group (AskAnalyzer; Gruber et al., 2015). First, we filtered the raw data between 11 and 16 Hz and then detected spindle events at frontal (F3, F4, Fz) and central (C3, C4, Cz) channels rereferenced to the average of mastoids. We used the criteria described in Schimicek et al. (1994). Only events with an amplitude $>12 \mu V$ and a duration between 500 and 2000 ms were considered. Further validation of the detected spindles was done using LDA in which the detected spindles were compared with a template that is generated based on the visual scoring of sleep spindles in 8730 min of PSG data from 189 healthy participants and 90 participants with sleep disorders. For our analyses, we only considered events that occurred during N2 and N3 sleep with an LDA score of 1.7 or higher (Anderer et al., 2010). Briefly, the algorithm was developed using the scoring of 12 PSG recordings by six independent experts. It incorporates information from central and occipital channels. First, the algorithm compares the absolute and relative power of nine frequency bands including theta, alpha, and high beta ($>16$ Hz) frequencies between a 3 s test window and a moving $10$ s baseline via a series of LDA separately for each channel. Second, the start and end of each event are determined by combining the posterior probabilities of all channels so that the number of microarousals per total sleep time is the same for both automatic and visual detections. Although EMG increases are not necessary for the identification of microarousals that occur during NREM sleep, some of the detected microarousals showed a concurrent increase in the amplitude of the EMG signal; however, this increase in EMG activity was not time locked to the high-frequency shifts of the EEG signal. The algorithm detects microarousals in all sleep stages; however, for the purpose of this study we selected microarousals that occurred during N2 and N3 only (Fig. 3, detected microarousals).

**Microarousal detection.** We detected microarousals semiautomatically using an algorithm developed by the SIESTA Group, which has been described in detail in Anderer et al. (2010). Briefly, the algorithm compares the absolute and relative power of nine frequency bands including theta, alpha, and high beta ($>16$ Hz) frequencies between a 3 s test window and a moving $10$ s baseline via a series of LDA separately for each channel. Second, the start and end of each event are determined by combining the posterior probabilities of all channels so that the number of microarousals per total sleep time is the same for both automatic and visual detections. Although EMG increases are not necessary for the identification of microarousals that occur during NREM sleep, some of the detected microarousals showed a concurrent increase in the amplitude of the EMG signal; however, this increase in EMG activity was not time locked to the high-frequency shifts of the EEG signal. The algorithm detects microarousals in all sleep stages; however, for the purpose of this study we selected microarousals that occurred during N2 and N3 only (Fig. 3, detected microarousals).

**The detection of transient microstates (micromeasures of alertness).** The detection of sleep microstates provides a more fine-grained scoring that detects transient changes in sleep architecture in 4 s epochs rather than 30 s, using an algorithm described in Jagannathan et al. (2018). The
algorithm uses the Hori scale (Tanaka et al., 1996) to classify the epochs into either awake, drowsy (N1), or N2 based on the signal from a subset of 14 electrodes distributed over frontal, central, parietal, occipital, and temporal regions. We extracted stimuli from N2 sleep only and removed stimuli with interstimulus intervals of \(4000\) ms. Then we filtered the data between 0.1 and 30 Hz before running the algorithm on an equal number of epochs in all conditions.

### EEG preprocessing and analyses

**Preprocessing.** We performed all the preprocessing steps in EEGLAB, version 14.1.1b (Delorme and Makeig, 2004). First, we excluded face and neck channels and downsampled the raw data from 183 EEG channels to 128 Hz. Then, we filtered the data between 0.1 and 40 Hz using a Butterworth bandpass filter. We performed bad channels rejection and interpolation as well as rereferencing to an average reference using the PREP pipeline described in Bigdely-Shamlo et al. (2015). Finally, we performed independent component analysis using the adaptive mixture independent component analysis toolbox and visually detected and discarded eye and muscle artefacts.

**Event-related analysis.** We epoched the preprocessed data into 3000 ms trials (\([-1000\) to 2000 ms relative to stimulus onset]). For each participant, we converted ERPs into percent power change relative to the 500 ms prestimulus-onset window using the formula \((\text{Data} - \text{mean baseline values})/\text{mean baseline values}\).

### Time-frequency analysis

Time-frequency representations (TFRs) were computed over 8000 ms epochs, \([-4000\) to 4000 ms relative to stimulus onset]). We choose relatively long epochs to avoid edge artifacts because of the transformation. We calculated TFRs by applying a 500 ms hanning window as taper on frequencies from 0.5 to 30 Hz in 0.5 Hz frequency steps and 5 ms temporal steps. Similar to ERPs, we converted participant-specific TFRs into percent power change relative to the 500 ms prestimulus-onset window.

### Intertrial phase coherence estimate

Following time-frequency transformation, we extracted the complex Fourier coefficient for each channel, frequency, and time point in every single trial. Then we computed the phase angles in each trial before finally averaging the single-trial intertrial phase coherence (ITPC) values over all trials per subject. We performed all analyses in FieldTrip (Oostenveld et al., 2011; https://fieldtriptoolbox.org).

### Statistical analyses

For all of our analyses, we randomly selected an equal number of events/epochs per condition from N2 and N3. Evoked events were defined as events that are detected by our algorithms in the 2000 ms poststimulus-onset window. Because of violations to the assumptions of parametric testing, we applied rank-based nonparametric tests using the nparLD function implemented in the nparLD package available in R (Noguchi et al., 2012). We report ANOVA-type statistics (ATS), \(p\) values \((\alpha = 0.05,\)
two sided), as well as effect sizes using relative treatment effects (RTE). Generally, RTEs represent the probability of the values from the whole dataset being smaller than a randomly chosen observation from the respective group. Therefore, RTE values range between zero and one. An RTE value of 0.5 means no effect. The higher the RTE value of one condition, the higher the probability that a randomly chosen value from that condition is larger than that randomly drawn from the whole dataset, and vice versa. When applicable, we performed post hoc tests via the nparLD function with Bonferroni’s correction for multiple comparisons. For repeated measures at different time points, we performed nonlinear mixed regression via generalized linear mixed models (GLMMs) implemented in the glmer function of the lme4 package in R (Bates et al., 2015). Both KC and microarousals were non-normally distributed. For K-complexes, we used a GLMM with a Poisson distribution. For microarousals, because of the presence of a notable amount of zero counts, we used the zero-inflated Poisson distribution implemented in the pscl package in R (Zeileis et al., 2008). We added our subjects as effects with random intercepts and slopes. We report the estimates of the fixed effects (B+) and their standard errors, z values, and p values. We performed post hoc interaction tests using marginal means estimates as implemented in the emmeans package in R with Tukey’s correction for multiple comparisons, and we report Cohen’s d effect sizes.

For the more temporally resolved analysis to compare between the latencies of the detected events, we binned the 2000 ms poststimulus intervals into bins of 100 ms, then we calculated the mean of the number of events in each bin for each subject and condition, before finally submitting these results to the permutation analysis in FieldTrip. The choice of the bin size is a compromise between a meaningful temporal resolution and a sufficient statistical power.

For ERP, TFR, and ITPC analyses, we selected equal numbers of epochs per condition (66.47 ± 56.25) for each subject. For this analysis, we averaged the signal from six frontal channels (F3, F4, F7, F8, Fcz, and Fz), using a time window from −500 to 2000 ms relative to stimulus onset. We calculated the grand average over all subjects in each condition before submitting the results to the nonparametric cluster-based permutation analysis in FieldTrip (Maris and Oostenveld, 2007). We performed two-sided paired-sample t tests followed by Monte Carlo’s approximation with 5000 permutations (cluster-alpha = 0.05 and critical alpha = 0.025). We report the sum of the t values (Σt), as well as Cohen’s d effect sizes calculated over all possible permutations, channels, time points, and frequencies in the cluster.

Results
Auditory stimulation influenced sleep microstructure but not macrostructure
First, we assessed the effects of auditory stimulation on sleep macrostructure. We found that auditory stimulation during sleep does not influence sleep macrostructure. That is, we found no change in sleep macrostructure from the adaptation to the experimental night (Fig. 4A). Specifically, we found an effect of Stage (ATS(2.61) = 62.75, p < 0.001, RTEwake = 0.24, RTEN1 = 0.26, RTEN2 = 0.79, RTEN3 = 0.70, RTEREM = 0.51), no effect of the Night (ATS(1) = 0.04, p = 0.84, RTENoNight = 0.48, RTEstim = 0.52), and no interaction Night × Stage (ATS(2.56) = 0.75, p = 0.51). Similarly, during the experimental night, we found no difference
in sleep macrostructure between periods of stimulation and periods of no stimulation (Fig. 4B). We found a main effect for Stage \((ATS(2.51) = 53.30, p < 0.001, \text{RTE}_{\text{wake}} = 0.28, \text{RTE}_{\text{N1}} = 0.18, \text{RTE}_{\text{N2}} = 0.82, \text{RTE}_{\text{REM}} = 0.74)\). There was no effect of the Stimulation \((ATS(1) = 3.36, p = 0.07, \text{RTE}_{\text{ADAPT}} = 0.5, \text{RTE}_{\text{EXP}} = 0.5)\), and no interaction Stimulation × Stage \((ATS(2.08) = 0.68, p = 0.52)\). We then opted for a more time-resolved analysis of sleep stages based on the Hori scoring system (Tanaka et al., 1996; Jagannathan et al., 2018), which uses 4 s epochs instead of the classical 30 s staging. Again, we found a main effect for Stage \((ATS(1.64) = 146.26, p < 0.001, \text{RTE}_{\text{wake}} = 0.28, \text{RTE}_{\text{N1}} = 0.18, \text{RTE}_{\text{N2}} = 0.82, \text{RTE}_{\text{REM}} = 0.74)\).
However, there was no effect of the Stimulation (ATS(1) = 3.36, \( p = 0.19 \), RTEADAPT = 0.5, RTEEXP = 0.5), but a significant interaction Stimulation × Stage (ATS(1.47) = 8.17, \( p = 0.02 \)). Post hoc pairwise test with Bonferroni’s correction for multiple comparisons revealed that auditory stimulation resulted in a higher number of sleep epochs (ATS(1) = 12.84, \( p < 0.001 \), RTEstim = 0.62, RTEnostim = 0.38) and a lower number of drowsy epochs (ATS(1) = 10.88, \( p < 0.001 \), RTEstim = 0.4, RTEnostim = 0.6), suggesting even deeper sleep during stimulation periods.

On the level of sleep microstructure, we compared the densities, that is, the numbers of events per minute of N2 and N3, of KCs, slow and fast spindles, as well as microarousals between stimulation and no-stimulation periods. We show that auditory stimulation significantly increased the densities of KCs (Fig. 4D; ATS(1) = 19.68, \( p < 0.001 \), RTEstim = 0.59, RTEnostim = 0.41) and spindles (Fig. 4E; slow: ATS(1) = 19.68, \( p < 0.001 \), RTEstim = 0.58, RTEnostim = 0.42 − fast: ATS(1) = 8, \( p = 0.005 \), RTEstim = 0.58, RTEnostim = 0.42). However, the increase in the density of microarousals did not reach statistical significance (Fig. 4F; ATS(1) = 20.16, \( p < 0.001 \), RTEstim = 0.55, RTEnostim = 0.45).

Next, we performed a more temporally resolved analysis of the auditory-induced changes in sleep microstructure. Specifically, we compared the numbers of KCs, spindles, and microarousals between Stimulus-ON and Stimulus-OFF periods during the experimental night. Stimulus-ON periods are 2000 ms poststimulus intervals (0–2000 ms relative to stimulus onset), whereas stimulus-OFF periods are 2000 ms intervals that start at least 2000 ms after the onset of the previous stimulus and during which no sounds are presented (Fig. 5A). We found significantly higher numbers of KCs (ATS(1) = 54.77, \( p < 0.001 \), RTEON = 0.68, RTEOFF = 0.32), slow spindles (ATS(1) = 9.13, \( p = 0.002 \), RTEON = 0.57, RTEOFF = 0.43), fast spindles (ATS(1) = 20.16, \( p < 0.001 \), RTEON = 0.55, RTEOFF = 0.45), and microarousals (ATS(1) = 8.45, \( p = 0.003 \), RTEON = 0.58, RTEOFF = 0.42) in the stimulus-ON than in the stimulus-OFF periods (Fig. 5B–E).
The brain responds selectively to unfamiliar voices during NREM sleep

We subsequently sought to investigate whether brain responses to auditory stimuli differ depending on their content, that is, the name and/or voice used in the stimulus. We used a nonparametric test from the nparLD package with two within factors, that is, Name (SON and UNs) and Voice (FV and UFV). KC responses to auditory stimuli showed a significant effect of Voice, as UFVs triggered more KCs than FVs (Fig. 6A; ATS(1) = 16.10, p < 0.001, RTEUFV = 0.76, RTEFV = 0.24), no effect of Name (ATS(1) = 0.09, p = 0.76, RTESON = 0.48, RTEUN = 0.52), and a significant interaction Name × Voice (ATS(1) = 11.86, p < 0.001). Post hoc tests revealed that the amount of KCs triggered by the combination FV-SON was marginally higher than that triggered by the combination FV-UNs (ATS(1) = 4.17, pBonf = 0.08, RTEFVSON = 0.56, RTEFVUN = 0.44), whereas there was no difference between UFV-SON and UFV-UNs (ATS(1) = 4.17, pBonf = 0.19, RTEUFVSON = 0.53, RTEUFVUN = 0.47).

For fast and slow spindles, however, there was no effect of Voice (Fig. 6B,C; slow: ATS(1,16) < 0.001, p = 0.99, RTEUFV = 0.5, RTEFV = 0.5 – fast: ATS(1,16) = 2.71, p = 0.10, RTEUFV = 0.53, RTEFV = 0.47), no effect of Name (slow: ATS(1,16) = 0.46, p = 0.5, RTESON = 0.51, RTEUN = 0.49 – fast: ATS(1,16) = 0.62, p = 0.43, RTESON = 0.49, RTEUN = 0.51), and no interaction Name × Voice (slow: ATS(1,16) = 0.11, p = 0.74, RTEFVSON = 0.51, RTEFVUN = 0.49, RTEUFVSON = 0.52, RTEUFVUN = 0.48 – fast: ATS(1,16) = 2.63, p = 0.10, RTEFVSON = 0.50, RTEFVUN = 0.55, RTEUFVSON = 0.47, RTEUFVUN = 0.47). Microarousals showed a main effect of Voice (Fig. 6D; ATS(1) = 9.14, p = 0.002, RTEUFV = 0.59, RTEFV = 0.41), no effect of Name (ATS(1) = 1.16, p = 0.29, RTESON = 0.53, RTEUN = 0.47) and no interaction Name × Voice (ATS(1) = 0.19, p = 0.66, RTEFVSON = 0.61, RTEFVUN = 0.58, RTEUFVSON = 0.44, RTEUFVUN = 0.37).

Figure 6E depicts the number of KCs and microarousals evoked by different stimulus types as compared with a 2000 ms no-stimulation intervals. Further, a temporally resolved analysis (see above, Statistical analyses) showed that the difference in the evoked KCs between FVs and UFVs occurred in the 100–800 ms poststimulus window (Fig. 6F; ∑t(16) = 25.35, p = 0.002, d = 1.02). For microarousals (Fig. 6G), the difference appeared 200 ms poststimulus onset, that is, 100 ms later than that of KCs, and was significant between 200 and 700 ms poststimulus (200–400 ms: t(16) = 5.92, p = 0.01, d = 0.84; 500–700 ms: t(16) = 5.19, p = 0.02, d = 0.76). It is worth mentioning that we observed a transient increase in the number of microarousals ~200 ms before the onset of UFV stimuli; however, this increase did not yield a significant statistical difference.

To illustrate whether the unfamiliarity of the voice is candidly the main cause of the difference in KC and microarousal responses, we divided the night into halves and hypothesized that brain responses to UFVs, but not to FVs, will decrease from the first to the second half, as the UFVs become more familiar with time. We modeled the change in the number of KCs from the first to the second half of the night using a GLMM with Poisson distribution fit by maximum likelihood (Fig. 7A). We found a main effect of Time (B^ = 1.17 ± 0.16, z = 7.51, p < 0.001), no effect of Time (B^ = −0.08 ± 0.11, z = −0.77, p = 0.44), and a significant interaction Time × Voice (z = −2.28, p = 0.02). Post hoc tests revealed a significant decrease of the UFV-triggered KCs from the first to the second half of the night (B^ = −0.31 ± 0.08, z = 3.54, p = 0.002, d = 0.31), whereas FV-evoked KCs did not change (B^ = −0.08 ± 0.11, z = 0.77, p = 0.47, d = 0.08). Conversely, for microarousals, a GLMM with zero-inflated Poisson distribution (Fig. 7B) demonstrated a marginally significant effect of Voice (B^ = 0.89 ± 0.51, z = 1.75, p = 0.08), no effect of Time (B^ = −0.19 ± 0.27, z = 0.73, p = 0.98), and no interaction Time × Voice (B^ = −0.09 ± 0.33, z = −0.27, p = 0.55). We provide a more detailed description of the GLMM results in Figure 7C,D.

UFVs evoke stronger K-complex-mediated brain responses during NREM sleep

Next, we aimed to examine the neural dynamics underlying the aforementioned differences in sleep microstructure. Therefore, we compared the ERPs between FVs and UFVs in the following conditions: (1) when the stimuli triggered KCs and (2) when no KCs were triggered. We found that when KCs were triggered, UFVs evoked a larger, more pronounced negative peak (Fig. 8A; ∑t(16) = −436.35, p > 0.001, d = 0.96) that resembles the N550 of the KC in its temporal and morphologic characteristics. Importantly, however, we observed that the peak of the N550 potential occurred later (~750–800 ms) than the usual time window (500–550 ms) and had a much smaller amplitude (20–50 μV) than that (~100 μV) previously reported in the literature (Colrain, 2005; Halász, 2005; Laurino et al., 2014). We illustrate that these discrepancies with the previous literature are because of the relatively large temporal window we defined for the detection of evoked KCs (2000 ms) as well as our decision to use an average reference compared with the mastoid reference used in earlier studies (Bastien and Campbell, 1992, 1994; Colrain, 2005; Halász, 2005; Laurino, 2014). Extended Data Figures 8-1 and 8-2 show the difference in the amplitude of the N550 according to the different referencing procedures as well as the latency jitter of the N550 peak on the single-trial level. Further, to confirm that
the difference in the ERPs does not reflect a difference in the amplitude of the evoked KCs, we compared the peak-to-peak amplitudes of the evoked KCs, as measured by the detection algorithm, between FVs and UFVs (Fig. 8B). Peak-to-peak amplitudes showed no significant effect of Voice (ATS(1) = 1.47, \( p = 0.23 \), RTEUFV = 0.48, RTEFV = 0.52), Name (ATS(1) = 0.71, \( p = 0.39 \), RTESON = 0.52, RTEUN = 0.48), and no interaction Voice × Name (ATS(1) = 1.47, \( p = 0.23 \), RTEUFVSON = 0.49, RTEUFVUN = 0.47, RTEFVSON = 0.54, RTEFVUN = 0.50). Intriguingly, when no KCs are evoked, we found no difference in the amplitude of the ERPs between UFVs and FVs (Fig. 8C; \( 0.3–0.37 \) s: \( \sum t(16) = 27.31 \), \( p = 0.13 \)).

**K-complex-mediated brain responses to UFVs reflect sensory processing**

We then speculated that the difference in the amplitude of the ERPs, if not because of a difference in the amplitude of the evoked KCs, might reflect more synchronized evoked responses to UFVs as compared with FVs. Therefore, we compared the phase consistency of brain responses between FVs and UFVs via the ITPC metric (Tallon-Baudry et al., 1996). We observed stronger ITPC following UFVs as compared with FVs in the delta band (1–4 Hz; Fig. 9A; \( \sum t(16) = 2413.98 \), \( p = 0.003 \), \( d = 1.06 \)) indicating more synchronized brain responses to UFVs. Figure 9B shows the ITPC contrast between FVs and UFVs indicating that difference in the ITPC is a product of stronger time-locked responses to UFVs. Similar to our ERP analysis, in the absence of the evoked KCs, ITPC values did not differ between FVs and UFVs (Fig. 9C; \( \sum t(16) = 2521.7 \), \( p = 0.001 \)).

Finally, we performed spectral analysis of brain responses to UFVs and FVs in the presence and absence of evoked KCs. We observed stronger low-frequency response (delta band; 1–4 Hz) to UFVs as compared with FVs (Fig. 9D,E) that started \( \sim 250 \) ms poststimulus onset. This delta response appeared independent of the presence (Fig. 9D; \( \sum t(16) = 9395.64 \), \( p < 0.001 \), \( d = 1.25 \)) or absence (Fig. 9E; \( \sum t(16) = 4097.56 \), \( p = 0.013 \), \( d = 0.86 \)) of the evoked KC.

Crucially, only in the presence of KCs, UFVs additionally elicited a significant increase in the power of higher frequencies (>16 Hz) as compared with FVs (Fig. 9D, \( \sum t(16) = 5020.62 \), \( p = 0.006 \), \( d = 1.04 \)) starting \( \sim 500 \) ms poststimulus.

**Discussion**

In this study, we presented sleepers with their own first names (SON) and two UNs spoken by either an FV or a UFV during a full night of sleep with polysomnography. We show that although auditory stimulation did change sleep architecture (Fig. 4A–C), it induced prominent, stimulus-specific changes in sleep microstructure. Generally, presenting auditory stimuli during NREM sleep increased the number of KCs, spindles, and microarousals (Fig. 5). UFVs triggered more KCs (Fig. 6A) and microarousals (Fig. 6D) than FVs. However, we found no difference in the amount of triggered KCs, spindles, or microarousals between SONs and UNs. The difference in the numbers of evoked KCs and microarousals between FVs and UFVs appeared \( \sim 100 \) ms poststimulus (200 ms for microarousals) and extended over the whole duration of the stimuli \( \sim 800 \) ms (Fig. 6F,G). Although the number of the UFV-triggered KCs decreased in the second half of the night, the numbers of evoked microarousals remained relatively stable throughout the night (Fig. 7). Moreover, in the presence of...
the auditory-evoked KC, UFVs triggered larger evoked responses (Fig. 8A), which did not reflect a difference in the amplitudes of the evoked KCs (Fig. 8B,C) but rather more synchronized brain responses to UFVs relative to FVs (Fig. 9A,B). Similarly, brain responses to UFVs demonstrated an increase in the power of high frequencies (>16 Hz), suggesting a stronger arousal reaction to UFVs (Fig. 9D). Crucially, we were not able to detect such differential brain responses between FVs and UFVs in the absence of the evoked KC.

It has previously been suggested that the more relevant the stimulus, the higher its tendency to trigger KCs (Halász, 2005). In this regard, our results pose UFVs as more relevant—or in evolutionary terms potentially more threatening (Blume et al., 2018)—and consequently more arousing to the sleeper than FVs. Indeed, the increase in microarousals following UFVs suggests a transient shift toward external processing of vital environmental stimuli. In the same vein, the decrease in the number of UFV-evoked KCs in the second half compared with the first half of the night (Fig. 7A) supports the notion that the sleeping brain continues to learn new information during sleep (Züst et al., 2019). That is, homeostatic regulatory processes alone cannot explain this observation because the number of FV-evoked KCs did not change, indicating a stimulus-specific attenuation of brain responses. It might be that the sleeping brain learns, through repeated processing, that an initially unexpected stimulus poses no immediate threat to the sleeper and consequently decreases its response to it. Conversely, in a safe sleep environment, the brain might be expecting to hear FVs and consistently inhibits any response to such stimuli to preserve sleep (Fig. 6E). Although this assumption remains speculative, it entails a thorough investigation of the ability of the brain to generate top-down predictions of the external sensory world during sleep. Nevertheless, our results suggest that the unfamiliarity of voice is a strong promoter of brain responses during NREM sleep.

What is the role of the auditory-evoked brain responses during NREM sleep? Central to such responses is the KC, that is, the most prominent sleep-specific response to sensory stimulation. To answer this question, we contrasted trials during which FVs and UFVs triggered KCs. In such trials, UFVs evoked a larger negative potential that resembles the N550 component of the KC (Fig. 7A). The N550 has been associated with large-scale neuronal silencing that protects sleep (Cash et al., 2009; Laurino et al., 2014) and, conversely, an arousal reaction that facilitates stimulus processing (Atienza et al., 2001). In our study, the amplitudes of the evoked KCs did not differ between FVs and UFVs, corroborating the previous literature presenting the evoked KC as an all-or-none phenomenon (Bastien and Campbell, 1992). Rather, the difference in the N550 amplitude between FVs and UFVs was because of more synchronized brain responses to UFVs as indicated by the ITPC values (Fig. 9A). Stimulus-induced phase modulations have been suggested to promote information processing and transmission in the cortex (Lakatos et al., 2013; Canavier, 2015; Voloh and Womelsdorf, 2016), and increased ITPC values have been associated with better cognitive performance (Hanslmayr et al., 2005; Eidelberg-Rothman et al., 2019) and enhanced attention (Joon Kim et al., 2007) during wakefulness. Together, our results suggest that the preferential brain responses to UFVs during NREM sleep reflect sensory processing. Importantly, the shorter poststimulus time window during which UFVs elicited KCs as compared with FVs, indicated by the narrower peak in Figure 6F, implies better temporal alignment of KCs following UFVs and emphasizes the contribution of the evoked KCs to the observed phase modulations. Finally, the stronger high-frequency responses (>16 Hz; Fig. 9) to UFVs that appeared only in the presence of KCs corroborate the role of the auditory-evoked KCs in promoting sensory processing of relevant information. Together, our findings suggest a central role for KCs in the extraction and processing of relevant external sensory information during NREM sleep.

We did not find any differences in brain responses between SON and UNs in contrast to the previous literature (Oswald et al., 1960; Perrin et al., 1999). One explanation might be that although sleep preserves low-level auditory processing, it attenuates higher order linguistic tracking (Makov et al., 2017). We speculate that this is because of the disruption of the activity of the large-scale networks necessary for higher order name processing (Démonet et al., 1992; Luke et al., 2002) because of the loss of long-range cortical connectivity during sleep (Massimini et al., 2005). Further research should elucidate the mechanisms of language tracking during sleep.

We show that the auditory-evoked spindles are not influenced by specific characteristics of the auditory stimuli (i.e., name or voice). However, the role of spindles in response to sensory stimuli during sleep is far from clear. Previous research has shown that spindles attenuate or even inhibit the processing of auditory information (De Gennaro and Ferrara, 2003; Schabus et al., 2012; Blume et al., 2018). Other work suggests that brain responses are preserved during spindles (Sela et al., 2016) and even argues for a role for spindles in the processing of memory-related sensory information (Anthony et al., 2018; Cairey et al., 2018; Jegou et al., 2019). More research should investigate the role of spindles in response to sensory information irrelevant to ongoing memory processes.

Finally, we found no relevant changes in sleep macrostructure and architecture because of auditory stimulation (Fig. 4). In fact, microarousals represent an integral part of healthy sleep as they ensure its reversibility (Halász et al., 2004). Hence, their slight increase should not be viewed as a disruption to normal sleep. Further, the analysis of microstates even indicates a shift toward deeper sleep in response to UVF stimuli (Fig. 4C), which is most likely a by-product of more auditory-evoked KCs that influenced the staging. However, the classical 30 s sleep staging might not be sensitive enough to capture subtle changes in sleep microstructure. Hence, the development and refinement of new fine-grained methods, such as the Hori-based microstate classification (Jagannathan et al., 2018) promises better monitoring of transient sleep fluctuations, especially in the presence of sensory perturbation.

In summary, sleep appears to be far from a homogenous state of unconsciousness. There are temporal windows in sleep during which the brain filters, extracts, and processes relevant external information. We speculate that such content-specific, dynamic reactivity to sensory signals enables the brain to enter a sentinel processing mode (Blume et al., 2018) during which it preserves the ability to efficiently engage in the important internal processes that are ongoing during sleep while remaining connected to the surrounding environment.

References
Ameen MS, Cheung LM, Hauser T, Hahn MA, Schabus M (2019) About the accuracy and problems of consumer devices in the assessment of sleep. Sensors 19:4160.
American Sleep Disorders Association (1992) EEG arousals: scoring rules and examples: a preliminary report from sleep disorders atlas task force of the American Sleep Disorders Association. Sleep 15:173–184.
Noguchi K, Gel YR, Brunner E, Konietschke F (2012) nparLD: an R software package for the nonparametric analysis of longitudinal data in factorial experiments. J Stat Soft 50:1–23.

Oostenveld R, Fries P, Maris E, Schoffelen J-M (2011) FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. Comput Intell Neurosci 2011:156869.

Oswald I, Taylor AM, Treisman M (1960) Discriminative responses to stimulation during human sleep. Brain 83:440–453.

Parapatics S, Anderer P, Gruber G, Saletu B, Saletu-Zyhlarz G, Dörfner G (2015) K-complex amplitude as a marker of sleep homeostasis in obstructive sleep apnea syndrome and healthy controls. Somnologie 19:22–29.

Perrin F, García-Larrea L, Mauguire F, Bastuji H (1999) A differential brain response to the subject’s own name persists during sleep. Clin Neurophysiol 110:2153–2164.

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.

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.

Rechtschaffen A, Kales A (1968) A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Bethesda, MD: U.S. Department of Health, Education, and Welfare.

Schabus M, Dang-Vu TT, Heib DPJ, Boly M, Desseilles M, Vandewalle G, Schmidt C, Albouy G, Darsaud A, Gais S, Degueldre C, Balteau E, Phillips C, Luxen A, Maquet P (2012) The fate of incoming stimuli during NREM sleep is determined by spindles and the phase of the slow oscillation. Front Neurol 3:40.

Schimwick P, Zeithofer J, Anderer P, Saletu B (1994) Automatic sleep-spike detection procedure: aspects of reliability and validity. Clin Electroencephalogr 25:26–29.

Schwarz JFA, Åkerstedt T, Lindberg E, Gruber G, Fischer H, Theorell-Haglöw J (2017) Age affects sleep microstructure more than sleep macrostructure. J Sleep Res 26:277–287.

Sela Y, Vyazovskiy VV, Cirelli C, Tononi G, Nir Y (2016) Responses in rat core auditory cortex are preserved during sleep spindle oscillations. Sleep 39:1069–1082.

Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J (1996) Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. J Neurosci 16:4240–4249.

Tanaka H, Hayashi M, Hori T (1996) Statistical features of hypnagogic EEG measured by a new scoring system. Sleep 19:731–738.

Voloh B, Womelsdorf T (2016) A role of phase-resetting in coordinating large scale neural networks during attention and goal-directed behavior. Front Syst Neurosci 10:18.

Zeileis A, Kleiber C, Jackman S (2008) Regression models for count data in R. J Stat Soft 27:1–25.

Züst MA, Ruch S, Wiest R, Henke K (2019) Implicit vocabulary learning during sleep is bound to slow-wave peaks. Curr Biol 29:541–553.e7.