

Embodiment of sleep-related words: Evidence from event-related potentials

Mareike J. Hülsemann  | Björn Rasch 

Division of Cognitive Biopsychology and Methods, Department of Psychology, University of Fribourg, Fribourg, Switzerland

Correspondence

Mareike J. Hülsemann and Björn Rasch, Division of Cognitive Biopsychology and Methods, Department of Psychology, University of Fribourg, Rue P.A. de Faucigny 2, 1701 Fribourg, Switzerland. Email: huelsemann@posteo.de (M. J. H.) and bjoern.rasch@unifr.ch (B. R.)

Funding information

H2020 European Research Council, Grant/Award Number: 677875

Abstract

Our thoughts can influence sleep, but the underlying mechanisms are unknown. According to the theory of “embodied cognition,” the semantic content of cognitive processes is represented by multimodal networks in the brain, which include body-related functions. Such multimodal representations could offer a mechanism, which explains mutual influences between cognition and sleep. Here we tested whether sleep-related words are represented in multimodal networks by examining the effect of congruent versus incongruent body positions on word processing during wakefulness. We experimentally manipulated the body position of 66 subjects (19–40 years old) between standing upright and lying down. Sleep- and activity-related words were presented around the individual speech recognition threshold. Our results show that word processing was facilitated in congruent body positions (sleep words: lying down and activity words: standing upright) compared with incongruent body positions, as indicated by a reduced N400 in the congruent condition with the lowest volume. In addition, early sensory components of the ERP (N180 and P280) were enhanced, suggesting that words were also acoustically better understood in a congruent body position. However, the difference in ERPs did not translate to differences on a behavioral level. Our results support the prediction of embodied processing of sleep- and activity-related words. Body position potentially induces a pre-activation of multimodal networks, thereby enhancing access to the semantic concepts of words related to current body position. The link between semantic meaning and body-related function could be a key element in explaining the influences of cognitive processing on sleep.

KEYWORDS

auditory word categorization, embodied cognition, language perception, multimodal representation, N400

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. *Psychophysiology* published by Wiley Periodicals LLC on behalf of Society for Psychophysiological Research

1 | INTRODUCTION

Our thoughts affect our sleep. Cognitive processes such as rumination and worries are key aspects underlying insomnia development and maintenance (Harvey et al., 2005; Riemann et al., 2010). Effective psychotherapeutic interventions to treat insomnia typically try to reduce, change or replace these types of “cognitive arousal.” In addition, “calming” of thoughts using diverse relaxation techniques can have sleep-promoting effects (Riemann et al., 2017). In particular, through a series of studies, we have shown that verbal suggestions to sleep deeper given before sleep can extend the time spent in objectively measured slow-wave sleep (SWS) (Cordi et al., 2014, 2015, 2020). However, the underlying mechanisms by which our thoughts or verbal suggestions influence neurophysiological sleep regulation are still unknown.

We have previously proposed that certain cognitive processes could remain active during sleep when initiated before sleep (Cordi et al., 2020). Thus, by initiating mental concepts related to “stress” or “sleep,” these concepts remain active during ongoing sleep and either impair or improve neurophysiological sleep regulation, respectively. Generally, the theoretical framework of “embodied cognition” offers an explanation for mutual influences between cognitive processes and body-related functions (Shapiro, 2010). According to this notion, mental processes involve simulations of body-related perceptions and actions, and therefore a close mutual link exists between cognitive processes and body-related sensorimotor systems of the brain (Gallese & Lakoff, 2005; Goldman, 1992). Evidence for embodied cognition is provided, for example, by brain imaging studies showing that concrete and abstract verbal concepts are capable of activating concept-related sensory, limbic, and sensorimotor areas in the brain: for example, the words “arm” versus “leg” activate the respective sensorimotor area (Boulenger et al., 2012). Conversely, higher compatibility between processed and performed action (action-compatibility effect; Glenberg & Kaschak, 2002; Zwaan et al., 2012) or a pre-activation of motor areas (via movement priming or neurostimulation) facilitates the processing of action words in classical language comprehension areas of the brain (e.g., Mollo et al., 2016; Pulvermüller et al., 2005). While there is abundant evidence supporting the embodiment of action-related concepts, the embodiment of sleep-related concepts has not yet been examined. Evidence for embodied processing of sleep-related concepts would provide an important aspect to take into consideration for possible mechanisms underlying sleep-promoting effects of verbal and suggestive relaxation techniques.

Here we conducted a first study examining whether the processing of sleep-related concepts is “embodied.” Similar to the “action-compatibility effect,” we examined whether the “sleep-compatibility” of our body posture would influence the

processing of sleep-related words. As sleep occurs almost exclusively in a relaxed lying body position, we expected a higher “sleep compatibility” for the processing of sleep-related concepts when participants were lying down and relaxing compared with an activated and upright position. Evidence suggests that body position per se influences neuronal activity and thus influencing cognitions and emotions as well (Thibault & Raz, 2016). Importantly, here we analyze whether the congruence between body state and processed word content modulates cognitive processes in an embodied way. In our study, participants listened to sleep- or activity-related words at four different volume levels close to their speech recognition threshold and had to categorize these words (see Figure 1a, for a summary of our procedure). To assess how difficult it was to process the meaning of the word, we focussed on event-related potentials (ERPs), with the N400 as our primary outcome parameter. The N400 amplitude in response to a stimulus is generally larger when the content of the stimulus is unexpected in a given context, thereby representing the ease of lexical access to the meaning of a word (Kutas & Federmeier, 2011; Lau et al., 2008). Therefore, we hypothesized that the N400 will be smaller when the processing of word categories is compatible with the body position (congruent condition: sleep words—lying down; activity words—standing upright) compared with an incompatible body position (incongruent condition: sleep words—standing upright; activity words—lying down). In addition, we hypothesized that on the behavioral level, categorization performance of the words would be better in the congruent compared with the incongruent condition. We expected the embodiment effects, both on the electrophysiological level as well as on a behavioral level, to generally be rather small and to primarily occur in processing conditions with a very high difficulty, that is, at the lowest volume levels of the presented words.

2 | METHOD

2.1 | Participants

The final sample size of 66 (50 female, 16 male) participants was comprised of both students from Fribourg university as well as young professionals. Mean age \pm standard deviation (*SD*) was 23 ± 4 years and ranged from 19 to 40 years. Initially, 71 subjects were invited to participate in the study. Four subjects were excluded due to a technical problem by which no responses were registered during the main task. One subject had to be excluded due to excessive sweating artifacts in the electroencephalogram (EEG).

Preconditions for eligibility were the following: no hearing damage (e.g., tinnitus), age between 18 and 40 years old, right-handedness, no acute health problems, no neurological or psychological impairments (e.g., depressive mood), no surgery within the past 3 months, no regular intake of medication

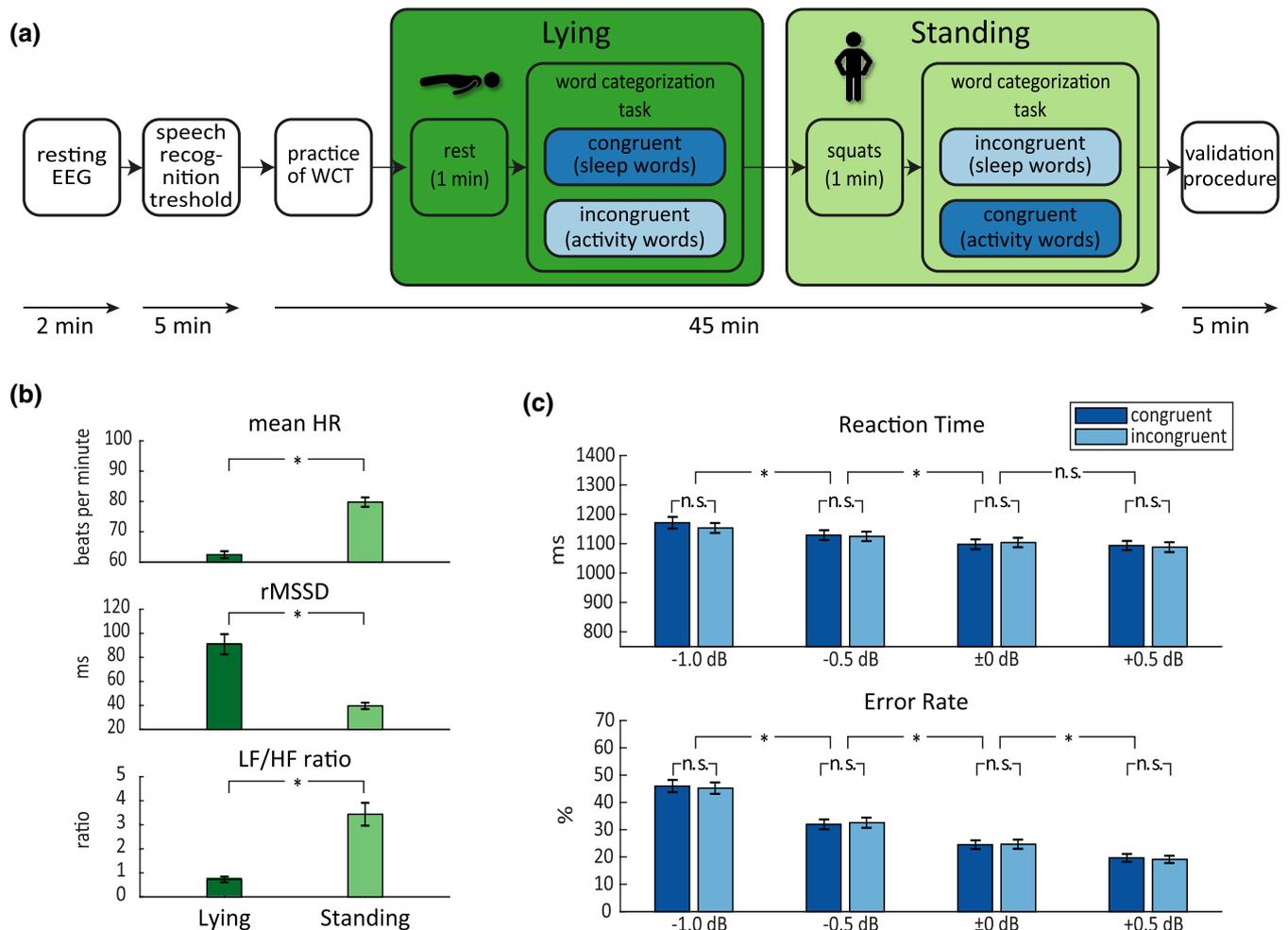


FIGURE 1 (a) Sequence of the experimental procedure. Subjects were seated in a sound attenuated and electrically shielded recording cabin. First, a two-minute resting EEG was recorded. Then the individual speech recognition threshold was acquired while participants were sitting. Participants practiced the word categorization task (WCT) while still sitting. They were then asked to lie down on a bed in the recording cabin and to breathe calmly and deeply for one minute. The first block of the word categorization task followed while participants were lying. Participants were then asked to stand up and do squats for one minute. The second block of the word categorization task followed while participants were standing. The order of the positions (lying vs. standing) was balanced across participants. Subsequently, in a validation procedure, stimulus material was presented visually and rated regarding category, valence, arousal and imaginability. This was performed while sitting. The cognitive tasks took roughly one hour, while the whole procedure including preparation and debriefing took two and half hours. (b) Effect of body position (lying vs. standing) on heart rate (HR) and heart rate variability (HRV) during the word categorization task. Mean HR (beats per minute, top) was reduced and rMSSD (root mean square of successive differences, middle) was increased during lying compared with standing. LF/HF ratio decreased during lying compared with standing, possibly indicating relatively more parasympathetic influence compared to sympathetic influences in the lying compared with the standing condition. (c) Mean reaction times and error rates according to stimulus onset in the word categorization task. In contrast to our hypotheses, congruency of the body position neither affected reaction times nor error rates. Generally, categorization performance was better for words presented at a higher compared with lower volumes, both for reaction times and error rates. Means \pm SE are indicated. * $p < .05$ for planned (b) and exploratory (c) pairwise comparisons

(e.g., insulin therapy for diabetes) except for oral contraception, no acute sleep disorders, at most moderate alcohol consumption according to World Health Organisation (WHO) guidelines, at most moderate tobacco use on maximally 4 days a week, and at most minor cannabis use on maximally 3 days a week. Women were neither pregnant nor breast-feeding.

Participants guaranteed completeness and accuracy of the information they provided and gave written informed consent prior to participation. The study was approved by the local

ethics committee and is in accordance with the Declaration of Helsinki (World Medical Association, 2001). Participants received either course credits or 30 Swiss francs for participation.

2.2 | Experimental procedure

The study took place in the sleep laboratory of the University of Fribourg. Subjects arrived either in the morning between

08h45 and 09h15 or in the afternoon between 16h30 and 17h15. Their eligibility to participate in the study was determined by questionnaires. Upon arrival, experimental procedures and information about the alleged aim of the study were provided. Participants were told that we wanted to examine the recognition performance of acoustic stimuli close to the speech recognition threshold and furthermore, that we wanted to examine whether body position had an additional influence on their recognition performance. However, the concept of embodiment was not mentioned. In particular, it was not mentioned that the auditory stimuli were related to the body positions.

Participants filled out trait questionnaires regarding their chronotype (Griefahn et al., 2001), sleep quality (PSQL, Buysse et al., 1989), extraversion (NEO-FFI extraversion scale, Borkenau & Ostendorf, 1993), stimulus reducing-augmenting (Vando RA-scale, Barnes, 1985), and absorption (TAS, Ritz & Dahme, 1995). These are of no further relevance for the present analyses.

Subsequently, they were seated in a sound attenuated and electrically shielded recording cabin and were prepared for simultaneous EEG, electrooculogram (EOG), and electrocardiogram (ECG) recordings. The experimenter, to guarantee a firm hold, attached ear clip headphones to both of the participant's ears.

Participants filled out a mood questionnaire (Steyer et al., 1997) before starting with the cognitive tasks. As this mood questionnaire contained a sleep-related word that was used later on in the main task, a second mood questionnaire was constructed with the same amount of activity-related words, used in the main task. For the cognitive tasks, all instructions were given via a computer screen or via headphones. However, the experimenter could be contacted at any time, in case of further questions. First, a two-minute resting EEG was recorded (one-minute eyes open and one-minute eyes closed in a randomized order across participants). Then the individual speech recognition threshold was acquired with a staircase procedure while participants were sitting (for details on the staircase procedure see Supporting Information).

Participants practiced the word categorization task while still sitting. Then subjects were informed that they had to execute the word categorization task in two positions, lying versus standing (Figure 1a). Order of positions was counterbalanced across subjects. They were informed that the difficulty of the task would be that some of the words would be presented very quietly. While these initial instructions were presented visually, all following instructions were then given via the headphones. First subjects were instructed to place themselves into the first position (lying down on a bed or standing upright), then subjects were informed that they should either breathe calmly and deeply (position lying) or do squats (position standing) for one minute. Breathing and doing squats were included in order to intensify a relaxed versus an activated

body state. Subjects could start the countdown of the minute individually and were informed when a minute had passed. Afterward, the word categorization task started. Subsequently, in a validation procedure, the words used in the word categorization task were presented visually and participants rated them according to category, valence, arousal, and imaginability without any time restriction (see Section 2.4.3 for results of the validation procedure). The validation was performed while sitting. Finally, participants filled out the original mood questionnaire for a second time and were then debriefed, in which they were informed about the embodiment hypothesis. The cognitive tasks took roughly one hour, while the whole procedure took one hour, two and half hours.

2.3 | Cognitive tasks

Cognitive tasks (speech recognition threshold, word categorization task, validation procedure) were presented on a 19" LCD monitor (HP Compaq LA 1956x) with 1,280 × 1,024 resolution and a 60 Hz refresh rate using E-Prime presentation software (Eprime 2.0, Psychological Software Tools, Pittsburgh, PA). The computer volume was set to 50% and the ear clip headphones (Philips SHS4700) were connected to the computer via an extension cable.

2.4 | Word categorization task

2.4.1 | Procedure

Participants heard spoken words belonging either to the category sleep or to the category activity. They had to decide which category the presented word belonged to. Subjects were informed that in some trials only noise would be presented. If they had heard noise or had not understood the word, they should not respond. The stimuli were presented at E-Prime volumes -500 , -250 , ± 0 , and $+250$ relative to the individual speech recognition threshold. These values correspond to roughly -1.0 dB SPL, -0.5 dB SPL, ± 0.0 dB SPL, and $+0.5$ dB SPL relative to the individual speech threshold.

The task consisted of one minute of deep breathing, 12 min categorizing while lying down, one minute of doing squats, and 12 min of categorizing while standing upright. A total of 480 trials were presented (240 while lying and 240 while standing). Nine unique sleep-related words, nine unique activity-related words (see Table 1), and two unique noise files (Brown Noise) were presented at four different volume levels (= 80 stimuli). The presentation of the 80 stimuli was repeated three times (= 240 stimuli). Stimuli were presented in random order within each of the three repetitions.

Each trial began with showing a fixation cross on the center of the screen for a variable time ranging between 100

TABLE 1 List of original, German stimulus material in the word categorization task and its translation to English

Sleep-related words		Activity-related words	
Dösen	<i>To doze</i>	Anfeuern	<i>To cheer so. on</i>
Kissen	<i>Pillow</i>	Eile	<i>Rush</i>
Schlapp ^a	<i>Groggily</i>	Hochleistung	<i>High performance</i>
Schnarchen	<i>To snore</i>	Hüpfen	<i>To jump</i>
Siesta	<i>Siesta</i>	Rumtoben	<i>To romp</i>
Trägheit	<i>Inertness</i>	Schnelligkeit	<i>Speed</i>
Träumen	<i>To dream</i>	Sportlich ^a	<i>Sporty</i>
Liegen	<i>To lie</i>	Stehen	<i>To stand</i>
Tiefschlaf	<i>Deep sleep</i>	Stress	<i>Stress</i>

Note: Words are alphabetically ordered.

^aThe words “groggily” and “sporty” were also presented in the mood questionnaires that were filled out by the participants before starting with the cognitive tasks.

and 500 ms. Then the audio file was presented. Each unique word stimulus had an onset delay of 101 to 132 ms; Brown noise had an onset delay of 30 ms. After the audio onset the fixation cross changed into a question mark. Subjects had a maximum of 2,400 ms to respond, starting from ca. 100 ms after stimulus onset. After the participants' response, a fixation cross was shown for 2,400 ms minus the reaction time. Therefore, subjects could not decrease the duration of the task by answering especially fast. The computer screen was not needed for executing the task as all instructions and relevant stimuli were presented via the headphones. However, participants had the possibility to see the computer screen when executing the task while standing, but could not see the screen while lying. This confounding variable is present in both—congruent and incongruent—conditions to the same extent, as both conditions include trials where participants were lying down and trials where participants were standing upright.

A number pad was used as a response device (Gembird USB NumberPad KPD-01). Participants were instructed to hold the number pad in their non-dominant hand and respond with the index finger of their dominant hand (all participants were right-handed). The left key “4” was always mapped to the first category and the right key “6” was always mapped to the last category. First and last relates to the order of mentioning the categories in the instructions. This order was counterbalanced across subjects. They were briefed to rest their index finger in between both response keys in between trials to guarantee accurate reaction time calculation. An animated figure was shown to visualize the response keys and accurate behavior. Subjects were instructed to respond as fast and as accurate as possible.

The stimuli were generated with a text-to-speech software provided by Microsoft Cognitive Services (<https://microsoft.com/cognitive-services/text-to-speech/#features>,

retrieved 2016/2017), using a male German (Austria) voice. Words were adjusted to have a 750 ms duration.

Participants first practiced the task while sitting. The practice contained a minimum of 12 and a maximum of 84 trials. It ended as soon as subjects responded in 80% of the ongoing trials correctly. During practice, participants received feedback on the correctness of their responses.

2.4.2 | Selection of stimulus material

Using the Edinburgh associative thesaurus (Kiss et al., 1973) a pool of words associated with the target terms “sleep” and “activity” was generated. All words that appeared in both lists, consisted of more than one word, were antonyms to the target term, had the same word stem as the target term, or were not easily translatable to German, were removed. Each list was sorted according to the strength of association with the target term. The first 42 words of each list (sleep and activity), together with 42 words related to the target term “family,” were presented to a total of $N = 134$ subjects (103 females and 31 males, 18–64 years old (median: 24)) in an online survey created using the platform SoSci Survey (Leiner, 2016). Subjects of the online survey rated all words of one list, intermixed with words of the two other lists. They rated valence, arousal, imaginability, and how strong they associated the words with the target term (target terms being sleep ($n = 50$), activity ($n = 36$), or family ($n = 48$)). Using the results of this survey, we calculated sensitivity (affiliation of each word to its corresponding list) and specificity (demarcation of each word from the other two lists). We then chose 18 words of each list, parallelizing them according to valence, imaginability, sensitivity, specificity, amount of syllables, and word frequency in German (Quasthoff et al., 2011). After piloting of the here presented experiment, the word pool was further reduced to nine words for each list, removing words that showed to be problematic in the pilot and adding four words for reasons of face validity (“to lie down,” “to stand,” “deep sleep,” and “high performance”). See Table 1 for the final word list and Table 2 for the stimuli characteristics.

2.4.3 | Validation of stimulus material

The stimuli used in the word categorization task were validated once more in the main experiment. All 18 stimuli from the word categorization task were visually presented in a randomized order while participants were sitting, after having performed the word categorization task. For each stimulus participants rated which category (sleep vs. activity) the word belonged to based on their personal preference. Additionally, they rated valence and arousal on a five-point SAM scale as

	Sleep words	Activity words	Statistics
Valence	3.45 ± 0.36	3.45 ± 0.42	$t(65) = 0.13, p = .90$
Arousal	1.90 ± 0.43	2.96 ± 0.73	$t(65) = 13.36, p < .01, \omega^2 = 0.57$
Imaginability	3.38 ± 0.42	3.32 ± 0.43	$t(65) = -1.66, p = .10$
Categorization error (%)	3.37 ± 7.54	1.18 ± 4.83	$t(65) = -2.85, p < .01, \omega^2 = 0.05$

Note: Categorization errors represent the words that were categorized into the opposite list as intended by the experimenters (e.g., categorizing a sleep-related word as activity-related). Valence and arousal were rated on a five-point SAM scale (valence: 1 = “negative,” 5 = “positive”; arousal: 1 = “calm,” 5 = “arousing”). Imaginability was rated on a four-point scale (1 = “not at all,” 4 = “very good”).

well as imaginability on a four-point scale from “not at all” to “very good.” There was no time restriction for responding and subjects could revise their responses. Sleep- and activity-related words did not differ in valence and imaginability. Activity-related words were much more arousing than sleep-related words. Sleep-related words were falsely categorized slightly more often compared to activity-related words. However, this effect was rather small and words were rarely miscategorized at all.

2.5 | Manipulation check: Effects of body position on HR and HRV

ECG was continuously recorded with silver-silver chloride (Ag/AgCl) electrodes (55 mm diameter, Kendall H98SG, Covidien). Two bipolar ECG channels were placed according to the standard lead II configuration. ECG was amplified by a BrainAmp ExG amplifier with an input impedance of 10 M Ω (Brain Products GmbH, Munich, Germany). Recordings, in AC mode, were sampled at 1,000 Hz with a 10 μ V resolution. The pass-band of the hardware filter was set to 0.016 to 250 Hz (see next section for further details).

Heart rate (HR) and heart rate variability (HRV) were analyzed using the software Kubios HRV Premium Version 3.2.0 (Tarvainen et al., 2014). Segments for both positions, consisting of 12 min continuous data each (recorded during the execution of the word categorization task) were processed separately. We used the built-in QRS detection algorithm and the automatic correction procedure to correct for artifacts and ectopic beats (see Kubios HRV User's Guide, Section 3.2, available at <http://kubios.uef.fi> for details). This automatic procedure leads to a median of 0.55% detected artifacts (range: 0.30%–11.65%). We exported mean HR (in beats per minute), root mean square of successive differences (rMSSD in ms), as well as the ratio between the power of the low frequency (LF) component (0.04–0.15 Hz) and the power of the high frequency (HF) component (0.15–0.40 Hz) of HRV (LF/HF ratio). Power values were calculated using the fast Fourier transform (FFT) procedure.

TABLE 2 Mean \pm SD of values for valence, arousal, imaginability, and categorization error for sleep- and activity-related words

2.6 | EEG recording and quantification

EEG was continuously recorded with an Acti-Cap electrode system (EasyCap GmbH, Herrsching, Germany) from 64 sites positioned according to the 10–10 electrode reference system (Chatrian et al., 1985). All sites of EEG were referenced online to FCz. AFz served as ground. Silver-silver chloride (Ag/AgCl) electrodes were used. EEG was amplified by a BrainAmp DC amplifier with an input impedance of 10 M Ω (Brain Products GmbH, Munich, Germany). Recordings, in AC mode, were sampled at 1,000 Hz with a 0.1 μ V resolution. Impedances of EEG electrodes were kept below 15 k Ω . The pass-band of the hardware filter was set to 0.016 to 250 Hz (high pass filter: first-order filter with -6 dB/octave; low pass filter: fifth-order Butterworth filter with -30 dB/octave). Recorded data were stored on a hard disk for later processing. EEG data were processed in MATLAB (R2018b, The MathWorks, Inc.), using the EEGLAB toolbox (v2019.1, Delorme & Makeig, 2004).

Raw EEG data were first pre-processed for running independent component analysis (ICA). ICA was used for removing eye, muscle, and heart artifacts. Data were high pass filtered (1 Hz) and low pass filtered (30 Hz) using a zero-phase Hamming-windowed sinc FIR filter implemented in EEGLAB (eegfiltnew.m; Widmann et al., 2015). Then the two blocks of the word categorization task were segmented and concatenated, excluding the one-minute breathing and one-minute squats. Noisy channels were rejected using the artifact subspace reconstruction (ASR) method (Chang et al., 2018; U.S. Patent Application No. 14/895440, 2014). Channels were rejected if they had flat line periods exceeding 5 s, were correlated less than 0.8 to their neighboring channels, or had line noise exceeding 4 SD. On average 3.74 (median: 2) channels were removed in each data set ranging from 0 to 18 removed channels (mainly Fp and AF channels, which were, therefore, not used for later analyses). EEG data were re-referenced to the average reference after noisy channel removal and the online reference channel FCz was re-used in the data. Cz was discarded after re-referencing to match the number of channels and the rank of data for ICA. Cz was chosen because its activity is best captured by neighboring electrodes and presumably the least

amount of information is lost by discarding this electrode. Data were temporally segmented into consecutive one-second epochs and artifactual epochs were rejected automatically using joint probability and kurtosis (Delorme et al., 2007). Rejection criteria for joint probability and kurtosis were set to 5 *SD* for single channels and 2 *SD* for all channels. Rejection of an epoch always led to the rejection of this epoch in all channels. This automatic procedure leads to an average of $14\% \pm 2\%$ of rejected epochs (range: 9%–18%). ICA was performed on the remaining segments, using the AMICA algorithm (Delorme et al., 2012). The resulting weight matrix was saved.

In a second step, raw EEG data were pre-processed for analysing ERPs. Raw EEG data were high pass filtered (0.1 Hz) and low pass filtered (30 Hz) using a zero-phase Hamming-windowed sinc FIR filter implemented in EEGLAB (Widmann et al., 2015). Then the two blocks of the word categorization task were segmented and concatenated, excluding the one-minute breathing and one-minute squats. Noisy channels, detected by the previous ASR procedure (described above) were removed, data were re-referenced to the average reference, the online reference channel FCz was re-used in the data, and Cz was discarded. The ICA weight matrix was imported and ICs were automatically classified using ICLabel (Pion-Tonachini et al., 2019). ICLabel provides the probability of each component to be generated by brain, muscle, eye, heart, line noise, channel noise, and others. We removed the ICs with highest probability of being generated by the eye (if probability exceeded 90%), the muscle (if probability exceeded 25%), and the heart (if probability exceeded 60%). Artifactual ICs were removed by setting their weights to 0 and re-calculating EEG activity with the resulting weight matrix. Thereafter rejected channels were interpolated (spherical). Trials were extracted by segmenting -200 to $+1,800$ ms around stimulus onset. Each trial was baseline corrected by subtracting the mean baseline activity (-200 to 0 ms according to stimulus onset) from all data points of the trial. Finally, artifactual epochs were rejected automatically using the same method and parameters as described above. This automatic procedure lead to an average of $17\% \pm 3\%$ rejected epochs (range: 7%–23%). Only trials with correct responses were kept for the ERP analysis. Furthermore, the first four presentations of each unique word stimuli were discarded (containing each volume level once), as these first presentations showed pronounced different brain responses compared to the stimuli that followed. Electrode positions Fp1, Fp2, AF7, AF3, AFz, AF4, AF8, O1, Oz, O2, and Iz were not included in the analyses because they were often noisy, and interpolation might have been poor due to their location.

Separate averages were computed for each individual, electrode, volume level (-1.0 dB, -0.5 dB, ± 0.0 dB, $+0.5$ dB), as well as for congruent (sleep words while lying and activity words while standing) and incongruent trials (sleep words while standing and activity words while lying). Amount of trials in the congruent versus incongruent conditions were

equalized individually for each subject and each volume level by choosing n random trials from the condition with the higher amount of trials (n being the amount of trials in the condition with lower amount of trials).

By visually inspecting the time course of the topography averaged over all participants and conditions, we detected three components in response to the speech stimuli. The exact temporal extent and maximum of each component were defined by inspecting the grand average ERP waveform. Suitability of the chosen time windows was checked and subsequently confirmed for each participant individually. Mean amplitude was calculated for each of the three components in the corresponding time windows (Luck, 2014), for each individual, electrode, volume level (-1.0 dB, -0.5 dB, ± 0.0 dB, $+0.5$ dB), as well as for congruent and incongruent trials.

2.7 | Statistical analyses

All statistical analyses were conducted with IBM Statistics for Windows version 26 (SPSS, Inc., IBM company) or MATLAB (R2018b, The MathWorks, Inc.), except otherwise specified. Significance levels were set to $p_{\text{two-tailed}} < .05$. Violations of sphericity were appropriately corrected by Greenhouse-Geisser ϵ_{GG} if $\epsilon_{\text{GG}} \leq 0.75$ (Box, 1954; Geisser & Greenhouse, 1958) or Huynh-Feldt ϵ_{HF} if $\epsilon_{\text{GG}} > 0.75$ (Huynh & Feldt, 1976). Post hoc t tests were conducted for further analyses of significant results. The measure of effect size ω^2 is reported for significant results (Hays, 1994). It is an estimator for the population effect Ω^2 , which specifies the systematic portion of variance in relation to the overall variance (Rasch et al., 2014a, 2014b). All power calculations were performed with G*Power (version 3.1.9.4, Faul et al., 2007) or MATLAB (R2018b). The sample was balanced according to starting time of the experimental session, order of positions, mapping of response keys, and sex. These factors were, therefore, controlled for and not included in statistical analyses. Data are always presented as mean \pm standard error (*SE*) except otherwise specified.

2.7.1 | Manipulation check: Effects of position on HR and HRV

Three dependent sample t tests were used to compare cardiovascular measures (heart rate, rMSSD, and LF/HF ratio) in the lying versus standing condition.

2.7.2 | Word categorization task—Behavioral data

Behavioral results (dependent variables: reaction time, error rate) in the word categorization task were analyzed by a 2×4

ANOVA with the within-subjects factors condition (congruent vs. incongruent) and volume (-1.0 dB SPL, -0.5 dB SPL, ± 0.0 dB SPL, $+0.5$ dB SPL relative to the individual speech recognition threshold) in a repeated measure design. For reaction time analysis only trials with correct responses were included. Furthermore, reaction time distributions for each subject were calculated. Reaction times that exceeded the criterion of three interquartile ranges above the third quartile of the individual distribution (Tukey, 1977) or that were faster than 200 ms, were deemed outliers and removed for statistical analyses.

2.7.3 | Word categorization task—ERP data

We conducted mass univariate analyses (Groppe et al., 2011) separately for the three detected ERP components (N180: 130–230 ms, P280: 230–330 ms, N400: 400–1,800 ms in fourteen 100 ms sub-epochs). Separate repeated measure ANOVAs were calculated for each electrode (54 electrodes included) and each 100 ms sub-epoch, including the factors condition (congruent vs. incongruent) and volume (-1.0 dB SPL, -0.5 dB SPL, ± 0.0 dB SPL, $+0.5$ dB SPL relative to the individual speech recognition threshold). In case of significant effects, we conducted post hoc *t* tests for the congruency by volume interaction (i.e., testing the congruency effect separately for each volume level) and the volume main effect. We corrected for multiple comparisons using cluster-based permutation tests (Groppe et al., 2011; Maris & Oostenveld, 2007). A cluster consists of electrodes connected in time and space that show significant effects. Only electrodes showing effects in the same direction were included in a cluster. The test reveals whether a cluster of a specific size is likely to be present in data where the condition labels of trials have been randomly shuffled. If the size of a cluster found in the original data is highly unlikely in permuted data, the cluster is classified as significant. Importantly, cluster-based permutation tests neither reveal the exact time period when significant differences are present nor the exact location of the detected effect (Sassenhagen & Draschkow, 2019). We included 10,000 permutations and used *F*-mass/*t*-mass (sum of all *F*/*t* values taken from electrodes showing significant effects (inferred from parametric statistics described above)) and cluster size (amount of significant electrodes for each time period) to quantify cluster significance.

3 | RESULTS

3.1 | Manipulation check: Effects of position on HR and HRV

We first tested whether our manipulation of body position (lying down vs. standing upright) during the performance of the

word-categorization task resulted in the expected differences in arousal markers. Consequently, HR and HRV measures were derived from the twelve minute segments, while participants executed the word categorization task. As expected in the lying position, subjects showed lower mean HR (beats per minute; $t(65) = -14.62, p < .01, \omega^2 = 0.62$), higher rMSSD (ms; $t(65) = 6.96, p < .01, \omega^2 = 0.26$), and lower LF/HF ratio ($t(65) = -2.71, p < .01, \omega^2 = 0.21$) compared to standing (Figure 1b). Manipulation of the body position, therefore, led to the expected differences in HR and HRV in the lying compared with the standing position. Thus, our manipulation of body position was successful and resulted in a calm versus activated body state for lying down and standing upright, respectively.

3.2 | Behavioral data

In the next step, we analyzed whether the difference in body positions affected performance in the detection success of the word-categorization task. According to our hypothesis of the embodiment of sleep- and activity-related words, we expected the categorization of sleep-related words to be better during a calm body state (i.e., lying down). In contrast, the categorization of activity-related words should profit from an activated body state (i.e., standing up). Thus, categorization performance should be generally better for words congruent to the body position (i.e., sleep-related words while lying down and activity-related words while standing up). In contrast, categorization should be generally worse for words incongruent to the body position. We expected the effect of congruency to be present in highly difficult perception conditions, that is, in our study predominantly when the volume was below the speech recognition threshold. Therefore, our main primary outcome was the interaction between the factor congruency (congruent vs. incongruent) and the factor volume (-1.0 dB SPL, -0.5 dB SPL, ± 0.0 dB SPL, $+0.5$ dB SPL).

In contrast to our hypothesis, we did not find a significant interaction between the factor congruency and volume, neither for reaction times ($F(3, 195) = 1.37, p = .26$) nor error rates ($F(1, 65) = 0.78, p = .51$). Furthermore, we also did not find a significant main effect of congruency (reaction times: $F(1, 65) = 1.37, p = .25$, error rates: $F(1, 65) = 0.08, p = .77$). The ANOVA had a minimum of 95% power to find effect sizes of $\Omega^2 \geq 0.01$ for the main effect congruency and the interaction congruency by volume, both for reaction times and error rates. We can safely exclude the existence of a small and larger effect of congruency between body position and word category on the categorization response to sleep- and activity-related words. Thus, the congruency between body position and word category did not influence the categorization speed nor categorization success, and this was not modulated by the volume of the presented words.

As expected, categorization performance was better for words presented at a higher volume compared with a lower

volume. The effect was present for both reaction times ($F(3, 195) = 42.03, p < .01, \omega^2 = 0.32$) and error rates ($F(3, 195) = 305.06, p < .01, \omega^2 = 0.78$). All volume levels differed significantly from each other with the exception of ± 0.0 dB and $+0.5$ dB with regard to reaction times (see Figure 1c, for post hoc tests).

3.3 | ERP data

3.3.1 | Characterization of the general event-related response to aurally presented words

We detected three components in response to the speech stimuli, which were presented at volumes around the individual speech recognition threshold (Figure 2). First, a centro-parietally distributed negative wave, ranging from 130 to 230 ms and peaking at 180 ms (N180). Second, a bilateral parieto-temporally distributed positive wave, ranging from 230 to 330 ms and peaking at 280 ms (P280). N180 and P280 were strongly and globally influenced by volume. The mean amplitude of both components increased with stimulus intensity (Figure 2b), as is typical for early sensory ERP components reflecting the physical properties of the stimuli (e.g., volume). Third, we detected a frontally distributed, slightly left oriented N400 peaking at 800 ms. This component emerged at 400 ms and slowly faded out. The N400 was reflected as a positively oriented wave in the parieto-occipital cortex. Moreover, we detected two distinct volume effects on the N400 (see Figure S5 in the Supporting Information). Initially, the N400 was less pronounced for the lowest volume level compared with all other volume levels. Latter, in contrast to the early exogenous components, the mean amplitude of the endogenous N400 component decreased with stimulus intensity: lower volumes led to a higher mean activity of the N400 compared with higher volumes. This indicates that it is more difficult, and it takes longer, to infer the meaning of a word presented at volumes below the speech recognition threshold. Descriptively the lowest volume always had the lowest mean activity for the N180 and P280, with increasing mean amplitude as volume increased. In contrast, for the N400, mean amplitude was always largest for the lowest volume and continuously decreased with increasing volume. Post hoc tests for the volume effect revealed that the mean amplitude of the N180 and P280 differed from each other for all volume levels, while for the N400 mean amplitude mainly differed for the lowest volume level (-1.0 dB SPL) compared with all other volume levels. Detailed analyses of the volume effect can be found in the Supporting Information.

3.3.2 | Congruency effect on the N400

We examined whether congruency of body position and word category led to the predicted stronger negativity of the

N400 for incongruent versus congruent trials. We expected the embodiment effect to be rather small and only to be present in highly difficult perception conditions. In our study, this is related to the volume levels below the speech recognition threshold. Because the N400 spanned across a large time window, from 400 to 1,800 ms after stimulus onset, and effects are not expected to be uniformly present over the complete period, we analyzed mean activity in 14 consecutive 100 ms epochs.

We found a significant cluster of electrodes, exhibiting a main effect of congruency for the N400 from 700 to 1,100 ms (F -mass: 73.30, $p < .05$; cluster size: 13, $p < .01$). Mean amplitude of the N400 component was larger, that is more negative, for incongruent than for congruent trials. The contributing electrodes were located frontally close to the midline. The effect was maximal at FCz in the time window 700–800 ms ($F(1, 65) = 7.87, p < .01, \omega^2 = 0.05$) and had the same direction for all significant electrodes and time windows. Effect sizes for the contributing electrodes ranged between $0.03 < \omega^2 < 0.05$, with an average ($\pm SD$) of $\omega^2 = 0.03 \pm 0.01$.

This main effect was further qualified by a congruency by volume interaction. A group of neighboring electrodes over the right fronto-central cortex showed significant congruency by volume interactions. These electrodes formed a significant cluster from 400 to 800 ms (F -mass: 56.02, $p < .01$; cluster size: 16, $p < .01$) and a second significant cluster from 1,200 to 1,500 ms (F -mass: 35.61, $p < .05$; cluster size: 10, $p < .05$). For the early cluster (400–800 ms) the effect was maximal at F2 in the time window 700–800 ms ($F(3, 195) = 4.95, p = .01, \omega^2 = 0.02$). Effect sizes for the contributing electrodes ranged between $0.01 < \omega^2 < 0.02$, with an average ($\pm SD$) of $\omega^2 = 0.01 \pm 0.00$. For the late cluster (1,200–1,500 ms) the effect was maximal at FC2 in the time window 1,200–1,300 ms ($F(3, 195) = 5.23, p = .01, \omega^2 = 0.02$). Effect sizes for the contributing electrodes ranged between $0.01 < \omega^2 < 0.02$, with an average ($\pm SD$) of $\omega^2 = 0.01 \pm 0.01$.

Additionally, a parieto-occipital group of electrodes exhibited significant congruency by volume interactions. These electrodes formed a significant cluster from 500 ms to 1,700 ms (F -mass: 130.80, $p < .01$; cluster size: 34, $p < .01$). The effect was maximal at P5 in the time window 900–1,000 ms ($F(3, 195) = 6.66, p < .01, \omega^2 = 0.03$). Effect sizes for the contributing electrodes ranged between $0.01 < \omega^2 < 0.03$, with an average ($\pm SD$) of $\omega^2 = 0.02 \pm 0.01$. However, this cluster should be interpreted cautiously, as it was mainly present due to a constant significant effect in the P5 electrode; none of the neighboring electrodes contribute consistently to the cluster.

To further examine this interaction effect and its direction, we conducted post hoc t tests. We tested the congruency effect separately for each volume level. Dependent sample t tests indicated a main effect of congruency for the lowest

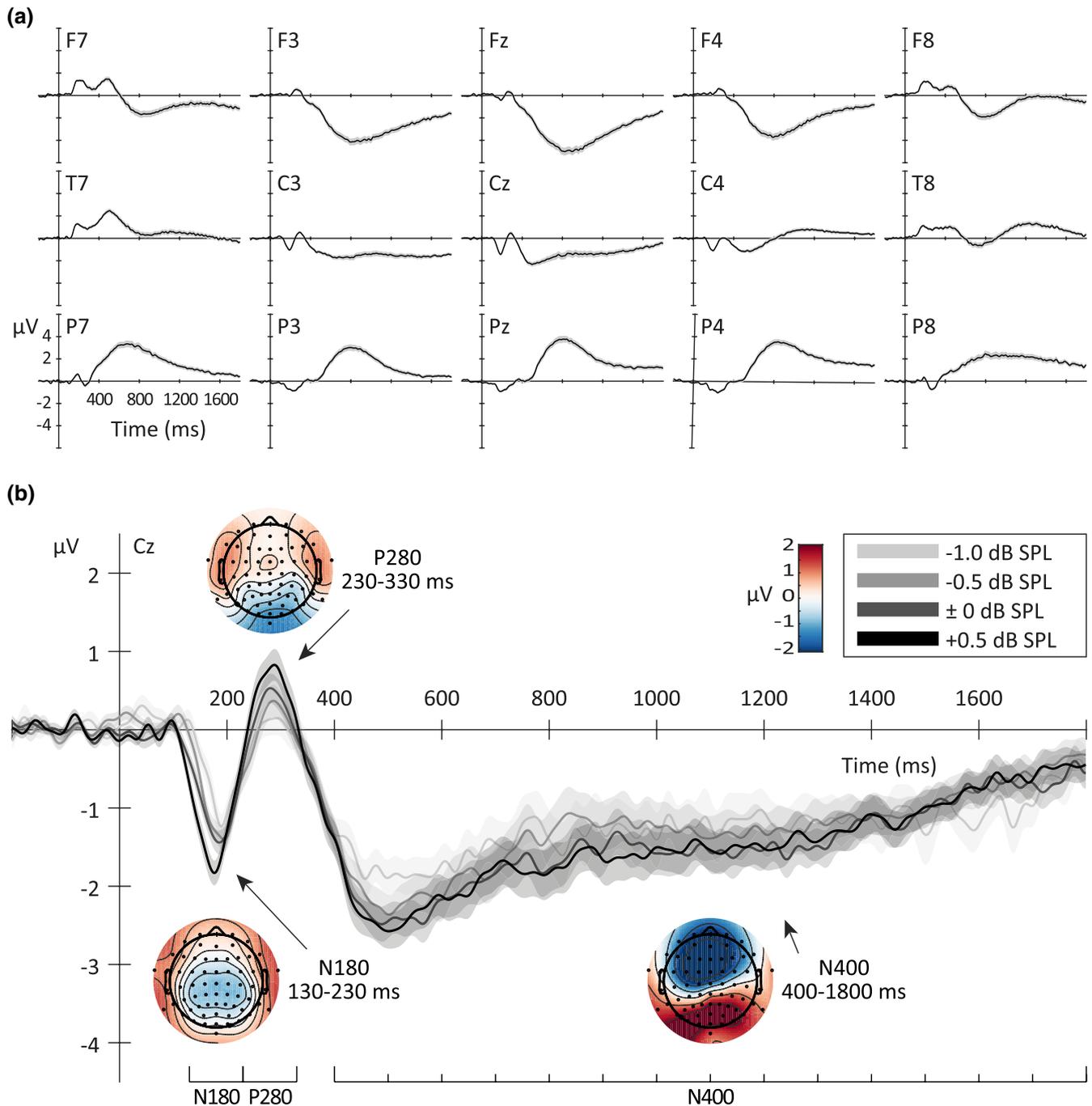


FIGURE 2 (a) Grand averages of the ERP (\pm SE) waveforms for a representative selection of electrodes averaged over all conditions and participants. The amplitude of N180 and P280 is relatively small compared to the N400 component which can be clearly detected as a negative going wave in frontal and central positions (upper two rows) and as a positive going wave in the parietal cortex (lower row). The N180 and P280 can be best detected in the central electrodes close to the midline. (b) Grand average of the ERP (\pm SE) at Cz, averaged over all participants and congruency conditions separated for the four different volume levels. The mean amplitude of the N180 and P280 increases with stimulus intensity (light grey—lowest volume to dark grey—highest volume). See Supporting Information for details on the volume effects. Additionally, the mean topographic maps over all subjects, conditions, and time windows used for analyses of the three components emerging in response to the presented speech stimuli in the word categorization task are depicted. The N180 is a centro-parietally distributed negative wave ranging from 130 to 230 ms and peaking at 180 ms. The P280 is a bilateral parieto-temporal distributed positive wave, ranging from 230 to 330 ms and peaking at 280 ms. The N400 (400–1,800 ms) is a frontally distributed, slightly left oriented negativity, reflected as positivity in the parietal cortex peaking at 800 ms (analyzed in 14 consecutive 100 ms epochs). All timing information relates to the stimulus onset at 0 ms

volume level (-1.0 dB SPL; Figure 3a). A group of neighboring electrodes at the right-central frontal cortex showed a stronger negativity for incongruent than for congruent

stimuli, forming a significant cluster (t -mass: 150.64, $p < .05$; cluster size: 56, $p < .05$). In this cluster of electrodes, the effect was present from 600 ms to 1,600 ms. Mean amplitude

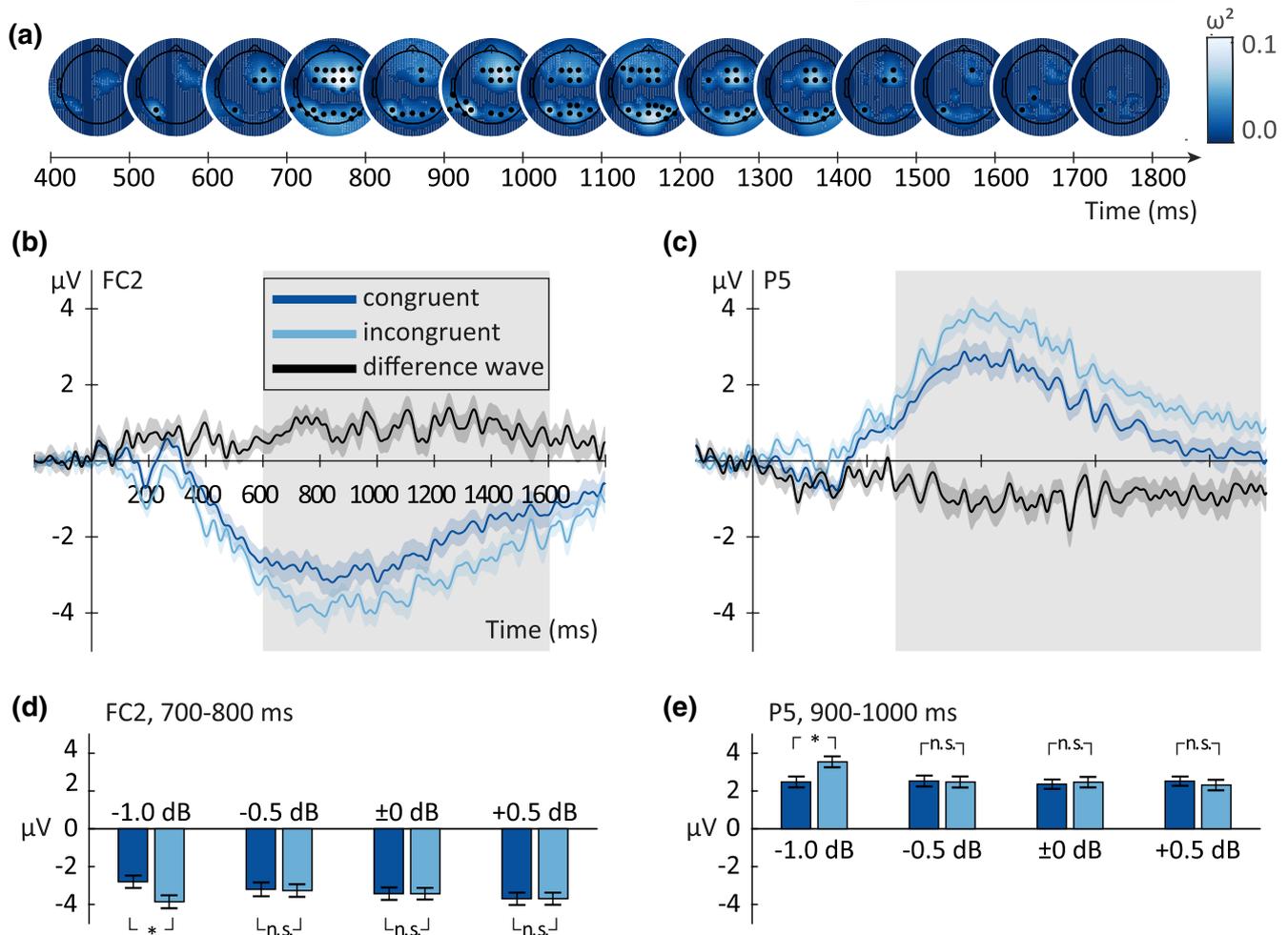


FIGURE 3 (a) Topographic maps of the effect sizes (ω^2) for the comparison of congruent versus incongruent trials at the lowest volume level (-1.0 dB SPL below the individual speech recognition threshold). Electrodes showing significant differences are plotted as black dots. The frontal cluster covers the right fronto-central cortex and emerges starting from 600 to 1,600 ms. The parietal cluster is present from 500 to 1,800 ms and spans both hemispheres. Lighter blue indicates larger effect sizes. Two exemplary electrodes from the frontal cluster (FC2, (b)) and from the parietal cluster (P5, (c)) showing the grand average ERP ($\pm SE$) for congruent (dark blue) and incongruent (light blue) trials, averaged over all participants for the lowest volume level (-1.0 dB SPL below the individual speech recognition threshold) are shown. Furthermore, the difference wave ($\pm SE$) between the two trial types is shown in black. The grey rectangle marks the period with significant differences between congruent and incongruent trials. The N400 amplitude is more pronounced for incongruent than for congruent trials, indicating that the processing of stimuli incongruent to the body position demands higher cognitive effort than the processing of stimuli congruent to the current body position. The bar plots show the mean N400 amplitude ($\pm SE$) for congruent (dark blue) and incongruent (light blue) trials averaged over the time window 700–800 ms for electrode FC2 (d) and averaged over the time window 900–1,000 ms for electrode P5 (e). Congruent and incongruent trials differ in amplitude only within the lowest volume level. The chosen electrodes and the chosen time windows reflect the periods where the effect was maximal for the frontal and parietal cluster, respectively. All electrodes contributing to a cluster had the same effect pattern as depicted in these exemplary electrodes. * $p < .05$

of the N400 component was larger, that is more negative, for incongruent than for congruent trials. The effect was maximal at FC2 in the time window 700–800 ms ($t(65) = 4.39$, $p < .01$, $\omega^2 = 0.12$) and had the same direction for all significant electrodes and time windows (Figure 3b). Effect sizes for the contributing electrodes ranged between $0.02 < \omega^2 < 0.12$, with an average ($\pm SD$) of $\omega^2 = 0.05 \pm 0.02$.

As for the interaction effect of the ANOVA, the post hoc t tests revealed an additional parieto-occipital cluster of electrodes

exhibiting significant congruency effects from 500 ms to 1,800 ms (t -mass: -159.04 , $p < .05$; cluster size: 62, $p < .05$; Figure 3a). The effect was maximal at P5 in the time window 900–1,000 ms ($t(65) = -4.45$, $p < .01$, $\omega^2 = 0.12$; Figure 3c). Effect sizes for the contributing electrodes ranged between $0.02 < \omega^2 < 0.12$, with an average ($\pm SD$) of $\omega^2 = 0.04 \pm 0.02$. This cluster mirrors the frontal cluster. The N400 is reflected as a positively oriented wave in the parietal cortex. Here we observe a stronger positivity for incongruent compared with congruent

trials. Therefore, as in the frontal cortex, incongruence leads to a larger mean amplitude of the N400. When analyzing the lowest volume level alone, this cluster is less dependent on a single electrode (P5, cf. cluster in the interaction effect). Therefore, we will carefully interpret the parieto-occipital cluster. It supports the interpretation of the right-frontal cluster of significant electrodes and could be a reflection of the frontal effect.

Such a congruency effect could not be detected in any of the other three volume levels (-0.5 dB SPL, ± 0.0 dB SPL, and $+0.5$ dB SPL; Figure 3d,e). Even though single electrodes showed significant congruency effects, none of these electrodes formed significant clusters and are, therefore, most likely chance findings. Across all t tests, we achieved a minimum of 95% power (calculated individually for each t test and electrode) to find effect sizes (Ω^2) of 0.05 ± 0.01 (range: 0.01–0.10) or larger. The ANOVA had a minimum of 95% power (calculated individually for each electrode) to find effect sizes (Ω^2) of 0.02 ± 0.01 (range: 0.00–0.05) or larger for the N400 main effect of congruency and 0.02 ± 0.01 (range: 0.01–0.04) or larger for the N400 congruency by volume interaction.

3.3.3 | Congruency effect on early components

To our knowledge, there are no reports of congruency effects on early ERP components. We, therefore, did not have specific hypotheses of congruency effects on the first two ERP components.

First, we examined the N180 component. We detected a significant cluster of electrodes at the parietal cortex showing a main effect of congruency for the N180 (F -mass: 58.49, $p < .05$; cluster size: 8, $p < .05$). These electrodes exhibited stronger negativity for congruent than for incongruent stimuli. The effect was only present in the left hemisphere. The effect was maximal at P1 ($F(1,65) = 10.57$, $p = .01$, $\omega^2 = 0.07$) and had the same direction for all significant electrodes. Effect sizes for the contributing electrodes ranged between $0.03 < \omega^2 < 0.07$, with an average ($\pm SD$) of $\omega^2 = 0.05 \pm 0.01$.

Furthermore, this group of neighboring electrodes showed a significant congruency by volume interaction, again forming a significant cluster (F -mass: 28.07, $p < .05$; cluster size: 7, $p < .05$). The effect was maximal at P5 ($F(3,195) = 5.35$, $p < .05$, $\omega^2 = 0.02$) and had the same direction for all significant electrodes. Effect sizes for the contributing electrodes ranged between $0.01 < \omega^2 < 0.02$, with an average ($\pm SD$) of $\omega^2 = 0.02 \pm 0.00$.

To further examine the N180 interaction effect and its direction, we conducted post hoc t tests. We tested the congruency effect separately for each volume level. Dependent sample t tests indicated a main effect of congruency for the lowest volume level (-1.0 dB SPL; Figure 4a). Similar to the effects of the ANOVA, a group of neighboring electrodes on the left parietal cortex showed a stronger negativity for

congruent than for incongruent stimuli, thereby forming a significant cluster (t -mass: -30.44 , $p < .01$; cluster size: 10, $p < .05$). The effect was maximal at P1 and P3 ($t(65) = -3.59$, $p < .01$, $\omega^2 = 0.08$) and had the same direction for all significant electrodes (Figure 4b). Effect sizes for the contributing electrodes ranged between $0.02 < \omega^2 < 0.08$, with an average ($\pm SD$) of $\omega^2 = 0.06 \pm 0.02$.

The effect tended to be reflected in a marginally significant cluster in the right frontal cortex (t -mass: 15.55, $p = .07$; cluster size: 7, $p < .05$). Effect sizes for the contributing electrodes ranged between $0.02 < \omega^2 < 0.04$, with an average ($\pm SD$) of $\omega^2 = 0.03 \pm 0.01$. Here we observe a stronger positivity for congruent compared to incongruent trials. Therefore, as in the parietal cortex, incongruence leads to a larger mean amplitude of the N180.

Such a congruency effect could not be detected in any of the other volume levels (-0.5 dB SPL, ± 0.0 dB SPL, and $+0.5$ dB SPL; Figure 4c). Across all t tests, we achieved a minimum of 95% power (calculated individually for each t test and electrode) to find effect sizes (Ω^2) of 0.06 ± 0.01 (range: 0.03–0.10) or larger. The ANOVA had a minimum of 95% power (calculated individually for each electrode) to find effect sizes (Ω^2) of 0.04 ± 0.01 (range: 0.03–0.05) or larger for the N180 main effect of congruency and of 0.03 ± 0.00 (range: 0.02–0.04) or larger for the N180 congruency by volume interaction.

Second, we examined the P280 component. A significant cluster of electrodes at the right frontal cortex showed a main effect of congruency for the P280 (F -mass: 80.42, $p < .05$; cluster size: 9, $p < .05$; Figure 4d). These electrodes exhibited a stronger positivity for congruent than for incongruent stimuli. The effect was maximal at FC4 ($F(1,65) = 13.24$, $p = .01$, $\omega^2 = 0.08$) and had the same direction for all significant electrodes (Figure 4e,f). Effect sizes for the contributing electrodes ranged between $0.02 < \omega^2 < 0.08$, with an average ($\pm SD$) of $\omega^2 = 0.06 \pm 0.02$. The effect was reflected as a negativity in the left parietal cortex; however, these electrodes did not form a significant cluster (F -mass: 42.94, $p = .06$; cluster size: 7, $p = .05$).

For the P280, there was no congruency by volume interaction at any electrode. The ANOVA had a minimum of 95% power (calculated individually for each electrode) to find effect sizes (Ω^2) of 0.03 ± 0.01 (range: 0.02–0.04) or larger for the P280 main effect of congruency and of 0.02 ± 0.00 (range: 0.02–0.03) or larger for the P280 congruency by volume interaction. We can, therefore, exclude the possibility of the presence of small and larger effects.

4 | DISCUSSION

Here we present evidence from ERPs sustaining that the processing of sleep- and activity-related words is embodied.

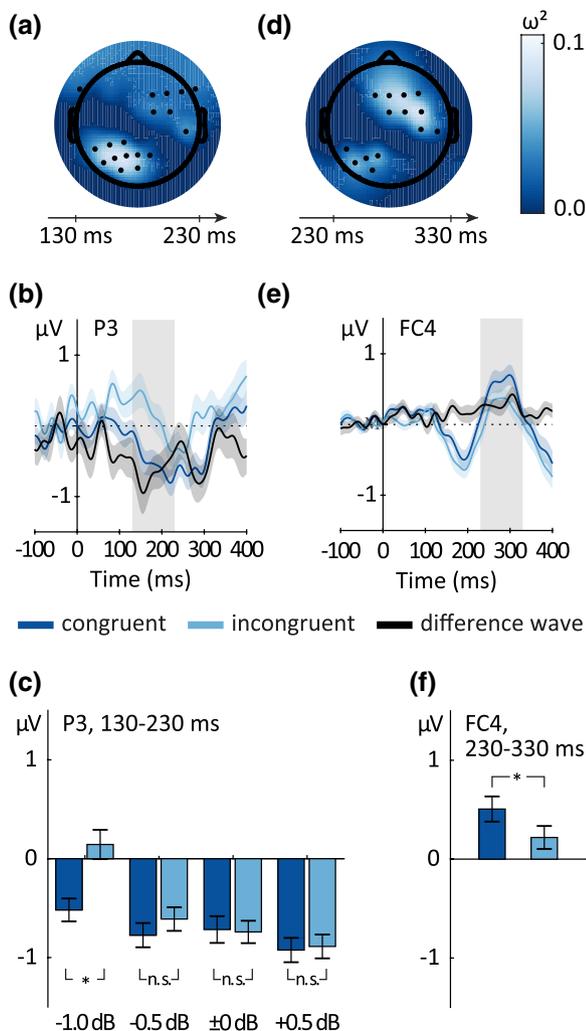


FIGURE 4 (a) Topographic map of the effect sizes (ω^2) for the comparison of congruent versus incongruent trials at the lowest volume level (-1.0 dB SPL below the individual speech recognition threshold) for the time window 130–230 ms post-stimulus onset. Electrodes showing significant effects are plotted as black dots. A significant cluster was formed by a group of electrodes in the left parietal cortex. Lighter blue indicates larger effect sizes. (b) Exemplary electrode (P3) from the parietal N180 cluster showing the congruency effect at the lowest volume level. The grand average ERP ($\pm SE$) for congruent (dark blue) and incongruent (light blue) trials is depicted as an average over all participants for the lowest volume level (-1.0 dB SPL below the individual speech recognition threshold). Furthermore, the difference wave ($\pm SE$) between the two trial types is shown in black. The N180 amplitude is more negative for congruent than for incongruent trials, indicating that speech stimuli congruent to the current body position are acoustically better understood than incongruent speech stimuli. The grey rectangle marks the period with significant differences between congruent and incongruent trials. (c) Mean N180 amplitude ($\pm SE$) for congruent (dark blue) and incongruent (light blue) trials averaged over the time window 130–230 ms for electrode P3. Only within the lowest volume level do congruent and incongruent trials differ in amplitude. (d) Topographic map of the effect sizes (ω^2) for the main effect congruency of the conducted ANOVA in the time window 230–330 ms post-stimulus onset. A significant cluster was formed by a group of electrodes in the right frontal cortex. (e) Exemplary electrode (FC4) from the frontal P280 cluster showing the grand average ERP ($\pm SE$) for congruent (dark blue) and incongruent (light blue) trials, averaged over all participants and all volume levels. Furthermore, the difference wave ($\pm SE$) between the two trial types is shown in black. The P280 amplitude is more positive for congruent than for incongruent trials. As for the N180, this could indicate that speech stimuli congruent to the current body position are acoustically better understood than incongruent speech stimuli. (f) Mean P280 amplitude ($\pm SE$) for congruent (dark blue) and incongruent (light blue) trials averaged over the time window 230–330 ms and all volume levels for electrode FC4. The difference between congruent and incongruent trials is present in all volume levels. The chosen electrodes are those where the effect was maximal for the N180 and P280, respectively. All electrodes contributing to a cluster had the same effect pattern as depicted in these exemplary electrodes. $*p < .05$

In particular, processing of words was facilitated in the congruent condition (sleep words—lying; activity words—standing) when words were presented at the lowest volume. This was indicated by a lower N400 amplitude in the congruent compared with the incongruent condition. As the N400 is typically associated with higher-order cognitive processing, we interpret the decrease in the N400 in the congruent condition as a facilitation of word processing caused by the compatibility of the body position under conditions where word processing is difficult. For earlier, more sensory ERP-components, the N180 and P280 showed larger deflections for congruent versus incongruent trials. Similarly, we observed larger deflections in these two early components with increased volume levels of the words. Thus, in addition to the facilitation of complex cognitive processing, a compatible body position might also improve sensory processing of words (e.g., the ease of understanding the word). We conclude that our ERP-results provide evidence for an embodiment of sleep- and activity-related words. In contrast to the electrophysiological level, we did not observe any effects of congruency between body position and word category on the behavioral level.

Our results of the embodiment of sleep-related words are in line with previous reports on the embodiment of action words. The action-compatibility effect predicts processing advantages, if processed and performed actions are compatible. In studies testing this effect, subjects are asked to judge the sensibility of sentences and respond with a movement. Response movements can be either congruent or incongruent to the movement described in the sentences. For example, Glenberg and Kaschak (2002) presented sentences implying a movement toward or away from the body. The response movement was significantly faster when its direction was compatible with the movement direction described in the sentence. Zwaan et al. (2012) presented sentences describing actions implying a forward or backward movement of the

whole body. Forward-sentences caused a congruent shift of the body. For sentences implying a backward movement, the detected forward shift was much less pronounced compared to sentences describing a forward shift. This was found even though subjects had to respond by left versus right movements to judge the sensibility of the sentences.

Furthermore, movement priming or neurostimulation facilitates the processing of action words in classical language comprehension areas of the brain. In Mollo et al. (2016) participants either pressed a button with their index finger or a pedal with their big toe. The authors found that arm-related and leg-related words led to stronger activity in the posterior superior temporal gyrus and hand motor cortex when the word was congruent to the response type than when it was incongruent. Similar to our study, their manipulation led to rather small neurophysiological congruency effects which did not affect the behavioral responses. Pulvermüller et al. (2005) demonstrated that transcranial magnetic stimulation (TMS) applied to the arm- and leg-motor cortex facilitates the processing of congruent words (arm- and leg-related action words, respectively) when contrasted to incongruent words; that is, responses to leg-related action words are faster compared with arm-related action words when the leg-motor cortex is stimulated 150 ms after visual word presentation.

These studies indicate that the pre-activation of motor networks facilitates the activation of compatible semantic networks. Conversely, semantic networks may be capable of activating corresponding motor networks. Evidence for the assumption of spreading network activation is for example provided by a study showing that listening to activity-related sentences can activate corresponding sensorimotor networks (Tettamanti et al., 2005). Two meta-analyses report that activated regions for action language processing and regions activated by motor execution, imagery, and observation are distinct, but overlap substantially (Courson & Tremblay, 2020; Yang & Shu, 2016). Therefore, these meta-analyses imply that language and action activate unique neuronal networks, which nevertheless overlap substantially and thereby are capable of activating each other.

In addition to movement priming, body positions have been successfully used as embodied “prime” in other studies, for example, in the context of emotions (Price & Harmon-Jones, 2015). In particular, activity from facial muscles has proven to modulate our experiences of emotions. A recent meta-analysis found supporting evidence for the hypothesis that facial feedback influences emotions (Coles et al., 2019). Importantly, these influencing effects are rather small and heterogeneous in their nature. Evidence for a strong effect of facial feedback on emotions (Strack et al., 1988) could not be replicated in a multi-site replication project (Wagenmakers et al., 2016), which fits well with our findings of rather small effect sizes and no effects of congruency on overt behavior.

Pulvermüller and colleagues have proposed that words are represented in cortex-wide semantic networks which reflect concepts that are grounded in perception and action (Kiefer & Pulvermüller, 2012; Pulvermüller, 1995). Furthermore, Damasio’s somatic marker hypothesis assumes numerous interactions between body states and cognition (Damasio, 1994, 1996). Similar to the studies on embodied language and emotion and based on these theoretical accounts we assume for our study that the body position of lying down induced a pre-activation of networks related to the semantic concepts of relaxation and sleep. Simultaneously, the processing of sleep-related words should activate the network representing the meaning of “sleep” but also bodily experiences, for example, the experience of lying down and being relaxed. Subsequently, corresponding words should be processed more efficiently than unrelated words.

The beneficial effect of a compatible body position on word processing occurred only at the lowest volume level for the N180 and the N400. The congruency effect for the P280 did not depend on volume. However, the differences between congruent and incongruent trials were descriptively larger for lower volumes compared to higher volumes. Generally, louder stimuli are easier to understand. Thereby, they quickly activate their corresponding semantic network. In contrast, in highly difficult perception conditions, for example, when playing the stimuli below the speech recognition threshold, the stimulus itself cannot quickly activate its corresponding semantic network. Therefore, if body position has pre-activated compatible semantic networks, this could help to process stimuli congruent to the current position. Consistent with this hypothesis, we mainly found embodiment effects at the lowest volume level. Furthermore, a compatible body position did not influence word processing on the behavioral level. Thus, the embodiment of semantic concepts might be of minor practical relevance for normal word processing during wakefulness.

Our result that a congruent body position facilitates the processing of sleep words opens up the possibility that the influence could also act in the opposite direction: the activation of sleep-related words might, in turn, facilitate the activation of networks related to sleep induction and sleep maintenance. Sleep and wakefulness are controlled by mutual inhibitory influences between the wakefulness-promoting arousal-system and the sleep-promoting network (Saper et al., 2005). A multimodal representation of sleep might include the sleep-promoting network. Thus, as our results show that sleep-related words are embodied, activation of these mental concepts might result in a more pronounced activation of the sleep-promoting network, thereby improving sleep quality. We assume in addition that the activity of mental concept can be initiated before sleep and remains active during sleep. Increased activity of a multimodal embodied representation of the concept “sleep” during sleep could be a promising

explanatory mechanism for the positive influence of pre-sleep cognitive intervention on later sleep processes (Cordi et al., 2014, 2015, 2020). Furthermore, we have recently provided evidence that the experimental activation of the mental concept of “sleep,” by presenting relaxation-related words during sleep, is, indeed, capable of extending the time spent in slow-wave sleep in young healthy participants (Beck et al., 2020). Thus, activation of multimodal representations of sleep-promoting networks is capable of influencing ongoing sleep processes. The study reported here adds to this notion by providing evidence that sleep-related concepts are in fact stored in a multimodal representation including body components, as body posture facilitates the processing of sleep-related words.

Our study has several limitations. First, we define the body position of “lying down” and a calm body state induced by deep breathing as congruent to sleep. We have seen that these are, indeed, key features associated with sleep (Saper et al., 2005). Nevertheless, our participants were fully awake and needed to be alert throughout the task. Not being actually asleep or active, but rather having induced a calm versus activated body state could explain why we only found small congruency effects on ERPs as well as our findings specifically related to the lowest volume level. Future studies should test the effects of the embodiment of sleep-related concepts using manipulations more closely related to sleep, for example, immediately waking up participants from the sleep state, comparing sleep deprived versus non-sleep deprived subjects or examining potential differences in the processing of sleep-versus activity-related words during the sleep state itself.

Second, we successfully manipulated physical arousal, which was manifested in a lower HR and higher HRV for lying down versus standing upright positions. However, one might need to differentiate between physical and cognitive arousal (Blum et al., 1967). These two types of arousal could exert very different effects on sleep and cognition. For example, cognitive arousal has shown to impair sleep (Wuyts et al., 2012). In contrast, physical arousal, in the form of exercise, does rather improve sleep (Kubitz et al., 1996; Stutz et al., 2019) and cognition (Chang et al., 2012). Thus, future studies might want to repeat the manipulation examining the effect of more or less cognitive arousal on the processing of sleep-related words.

Third, the here elicited auditory N400 component deviates in latency and topography from the classic, visual N400 component. It is more frontally distributed (in contrast to the classic posterior topography) and rather late (peaking at 800 ms post-stimulus) compared to the classical N400 literature. These differences are potentially caused by stimulus modality (auditory vs. visual) and type of incongruence: typically incongruence is studied as semantic incongruence within sentence settings. Here we have independent single words, presented auditorily, where incongruence is

manipulated due to body position. However, the N400 is best characterized by a common functionality, but not so much by its localization, time course or a specific mental operation (Kutas & Federmeier, 2011, p. 623). Furthermore, the literature suggests that the N400 component is not generated in a single structure but by various widely distributed cortical networks (Kutas & Federmeier, 2011; Lau et al., 2008) which could explain differences in topography depending on the study design.

Fourth, the time-window of the N400 does overlap with the time-window where participants overtly responded with a button press. Therefore, we cannot exclude the possibility that the ERPs overlap with the movement-related cortical potentials. However, as each trial included in the analyses contains a button press, the differences in the ERPs are unlikely to be caused by participants' responses. In addition, the N400 peak (800 ms post-stimulus onset) considerably precedes the mean reaction time ($1,120 \pm 16$ ms post-stimulus onset).

To conclude, our results support the prediction of the embodiment of sleep- and activity-related words. The effects of body position on the mental processing of word meanings are small and do not affect behavior measured in the form of reaction times and error rates. Future studies need to replicate our findings and further test the theoretical consequences of the embodiment of sleep-related words regarding current notions on the influence cognition has on sleep.

ACKNOWLEDGMENTS

This study was supported by a grant from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement MemoSleep No. 677875). The funding institution had no further role in the study design, the collection, analysis, and interpretation of data, the writing of the manuscript, and the decision to submit the paper for publication. We thank Ewald Naumann and Angelika Dierolf for advice regarding the event-related potential analyses and helpful feedback on earlier versions of this manuscript. We would like to thank Louisa Clarke for helpful comments on this manuscript. We furthermore thank our students for assistance in data collection and all subjects for their participation.

CONFLICT OF INTEREST

No conflicts of interest were associated with this research.

AUTHOR CONTRIBUTIONS

Mareike J. Hülsemann: Data curation; Formal analysis; Investigation; Methodology; Project administration; Software; Validation; Visualization; Writing-original draft; Writing-review & editing. **Björn Rasch:** Conceptualization; Funding acquisition; Methodology; Resources; Supervision; Validation; Visualization; Writing-original draft; Writing-review & editing.

ORCID

Mareike J. Hülsemann  <https://orcid.org/0000-0001-5542-938X>

Björn Rasch  <https://orcid.org/0000-0001-7607-3415>

REFERENCES

- Barnes, G. E. (1985). The Vando R-A scale as a measure of stimulus reducing-augmenting. In J. Strelau (Ed.), *The series in clinical and community psychology* (pp. 171–180). Hemisphere Publishing Corp/Harper & Row Publishers.
- Beck, J., Loretz, E., & Rasch, B. (2020). Exposure to relaxing words during sleep promotes slow-wave sleep and subjective sleep quality. *BioRxiv*. Advance online publication. <https://doi.org/10.1101/2020.12.16.423012>
- Blum, G. S., Geiwitz, P. J., & Stewart, C. G. (1967). Cognitive arousal: The evolution of a model. *Journal of Personality and Social Psychology*, 5(2), 138–151. <https://doi.org/10.1037/h0024069>
- Borkenau, P., & Ostendorf, F. (1993). *NEO-Fünf-Faktoren Inventar (NEO-FFI) nach Costa und McCrae*. Hogrefe.
- Boulenger, V., Shtyrov, Y., & Pulvermüller, F. (2012). When do you grasp the idea? MEG evidence for instantaneous idiom understanding. *NeuroImage*, 59(4), 3502–3513. <https://doi.org/10.1016/j.neuroimage.2011.11.011>
- Box, G. E. P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification. *The Annals of Mathematical Statistics*, 25(2), 290–302. <https://doi.org/10.1214/aoms/117728786>
- Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research*, 28(2), 193–213. [https://doi.org/10.1016/0165-1781\(89\)90047-4](https://doi.org/10.1016/0165-1781(89)90047-4)
- Chang, C.-Y., Hsu, S.-H., Pion-Tonachini, L., & Jung, T.-P. (2018). *Evaluation of artifact subspace reconstruction for automatic EEG artifact removal*. 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI (pp. 1242–1245). <https://doi.org/10.1109/EMBC.2018.8512547>
- Chang, Y. K., Labban, J. D., Gapin, J. I., & Etnier, J. L. (2012). The effects of acute exercise on cognitive performance: A meta-analysis. *Brain Research*, 1453, 87–101. <https://doi.org/10.1016/j.brainres.2012.02.068>
- Chatrian, G. E., Lettich, E., & Nelson, P. L. (1985). Ten percent electrode system for topographic studies of spontaneous and evoked EEG activities. *American Journal of EEG Technology*, 25(2), 83–92. <https://doi.org/10.1080/00029238.1985.11080163>
- Coles, N. A., Larsen, J. T., & Lench, H. C. (2019). A meta-analysis of the facial feedback literature: Effects of facial feedback on emotional experience are small and variable. *Psychological Bulletin*, 145(6), 610–651. <https://doi.org/10.1037/bul0000194>
- Cordi, M. J., Hirsiger, S., Méritat, S., & Rasch, B. (2015). Improving sleep and cognition by hypnotic suggestion in the elderly. *Neuropsychologia*, 69, 176–182. <https://doi.org/10.1016/j.neuropsychologia.2015.02.001>
- Cordi, M. J., Rossier, L., & Rasch, B. (2020). Hypnotic suggestions given before nighttime sleep extend slow-wave sleep as compared to a control text in highly hypnotizable subjects. *International Journal of Clinical and Experimental Hypnosis*, 68(1), 105–129. <https://doi.org/10.1080/00207144.2020.1687260>
- Cordi, M. J., Schlarb, A. A., & Rasch, B. (2014). Deepening sleep by hypnotic suggestion. *Sleep*, 37(6), 1143–1152. <https://doi.org/10.5665/sleep.3778>
- Courson, M., & Tremblay, P. (2020). Neural correlates of manual action language: Comparative review, ALE meta-analysis and ROI meta-analysis. *Neuroscience and Biobehavioral Reviews*, 116, 221–238. <https://doi.org/10.1016/j.neubiorev.2020.06.025>
- Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. G.P. Putnam.
- Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 351(1346), 1413–1420. <https://doi.org/10.1098/rstb.1996.0125>
- Delorme, A., & Makeig, S. (2004). Eeglab: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Delorme, A., Palmer, J., Onton, J., Oostenveld, R., & Makeig, S. (2012). Independent EEG sources are dipolar. *PLoS One*, 7(2), e30135. <https://doi.org/10.1371/journal.pone.0030135>
- Delorme, A., Sejnowski, T., & Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *NeuroImage*, 34(4), 1443–1449. <https://doi.org/10.1016/j.neuroimage.2006.11.004>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/bf03193146>
- Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, 22(3), 455–479. <https://doi.org/10.1080/02643290442000310>
- Geisser, S., & Greenhouse, S. W. (1958). An extension of box's results on the use of the F distribution in multivariate analysis. *The Annals of Mathematical Statistics*, 29(3), 885–891. <https://doi.org/10.1214/aoms/1177706545>
- Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, 9(3), 558–565. <https://doi.org/10.3758/BF03196313>
- Goldman, A. I. (1992). In defense of the simulation theory. *Mind & Language*, 7(1–2), 104–119. <https://doi.org/10.1111/j.1468-0017.1992.tb00200.x>
- Griefahn, B., Kunemund, C., Brode, P., & Mehnert, P. (2001). Zur Validität der deutschen Übersetzung des Morningness-Eveningness-Questionnaires von Horne und Ostberg. The Validity of a German Version of the Morningness-Eveningness-Questionnaire Developed by Horne and Ostberg. *Somnologie*, 5(2), 71–80. <https://doi.org/10.1046/j.1439-054X.2001.01149.x>
- Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, 48(12), 1711–1725. <https://doi.org/10.1111/j.1469-8986.2011.01273.x>
- Harvey, A. G., Tang, N. K. Y., & Browning, L. (2005). Cognitive approaches to insomnia. *Clinical Psychology Review*, 25(5), 593–611. <https://doi.org/10.1016/j.cpr.2005.04.005>
- Hays, W. L. (1994). *Statistics* (5th ed.). Harcourt Brace College Publishers.
- Huynh, H., & Feldt, L. S. (1976). Estimation of the box correction for degrees of freedom from sample data in randomized block and

- split-plot designs. *Journal of Educational Statistics*, *1*(1), 69–82. <https://doi.org/10.2307/1164736>
- Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *48*(7), 805–825. <https://doi.org/10.1016/j.cortex.2011.04.006>
- Kiss, G. R., Armstrong, C., Milroy, R., & Piper, J. (1973). An associative thesaurus of English and its computer analysis. In A. J. Aitken, R. W. Bailey, & N. Hamilton-Smith (Eds.), *The computer and literary studies* (pp. 153–165). Edinburgh University Press.
- Kothe, C. A. E., & Jung, T.-P. (2014). *U.S. Patent Application No. 14/895,440*.
- Kubitz, K. A., Landers, D. M., Petruzzello, S. J., & Han, M. (1996). The effects of acute and chronic exercise on sleep. A meta-analytic review. *Sports Medicine*, *21*(4), 277–291. <https://doi.org/10.2165/00007256-199621040-00004>
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–647. <https://doi.org/10.1146/annurev.psych.093008.131123>
- Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. *Nature Reviews Neuroscience*, *9*(12), 920–933. <https://doi.org/10.1038/nrn2532>
- Leiner, D. J. (2016). *SoSci survey* (Version 2.6.00-i) [Computer Software]. <https://www.sosicisurvey.de>
- Luck, S. J. (2014). *An introduction to the event-related potential technique. A Bradford book* (2nd ed.). The MIT Press.
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>
- Mollo, G., Pulvermüller, F., & Hauk, O. (2016). Movement priming of EEG/MEG brain responses for action-words characterizes the link between language and action. *Cortex*, *74*, 262–276. <https://doi.org/10.1016/j.cortex.2015.10.021>
- Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, *198*, 181–197. <https://doi.org/10.1016/j.neuroimage.2019.05.026>
- Price, T. F., & Harmon-Jones, E. (2015). Embodied emotion: The influence of manipulated facial and bodily states on emotive responses. *Wiley Interdisciplinary Reviews. Cognitive Science*, *6*(6), 461–473. <https://doi.org/10.1002/wcs.1370>
- Pulvermüller, F. (1995). Neurobiologie der Wortverarbeitung. *Naturwissenschaften*, *82*(6), 279–287. <https://doi.org/10.1007/BF01134525>
- Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience*, *21*(3), 793–797. <https://doi.org/10.1111/j.1460-9568.2005.03900.x>
- Quasthoff, U., Fiedler, S., & Hallsteinsdóttir, E. (Eds.). (2011). *Frequency dictionaries: Vol. 1. Frequency dictionary German: Deu = Häufigkeitwörterbuch Deutsch* (1. Aufl.). : Leipziger Univ.-Verl.
- Rasch, B., Friese, M., Hofmann, W. J., & Naumann, E. (2014a). *Quantitative Methoden 1: Einführung in die Statistik für Psychologen und Sozialwissenschaftler* (4., überarb. Aufl.). Springer-Lehrbuch.
- Rasch, B., Friese, M., Hofmann, W. J., & Naumann, E. (2014b). *Quantitative Methoden 2: Einführung in die Statistik für Psychologen und Sozialwissenschaftler* (4., überarb. Aufl.). Springer-Lehrbuch.
- Riemann, D., Baglioni, C., Bassetti, C., Bjorvatn, B., Dolenc Groselj, L., Ellis, J. G., Espie, C. A., Garcia-Borreguero, D., Gjerstad, M., Gonçalves, M., Hertenstein, E., Jansson-Fröjmark, M., Jennum, P. J., Leger, D., Nissen, C., Parrino, L., Paunio, T., Pevernagie, D., Verbraecken, J., ... Spiegelhalder, K. (2017). European guideline for the diagnosis and treatment of insomnia. *Journal of Sleep Research*, *26*(6), 675–700. <https://doi.org/10.1111/jsr.12594>
- Riemann, D., Spiegelhalder, K., Feige, B., Voderholzer, U., Berger, M., Perlis, M., & Nissen, C. (2010). The hyperarousal model of insomnia: A review of the concept and its evidence. *Sleep Medicine Reviews*, *14*(1), 19–31. <https://doi.org/10.1016/j.smrv.2009.04.002>
- Ritz, T., & Dahme, B. (1995). The absorption scale-basic concept, psychometric properties, and dimensions of a German adaptation. *Diagnostica*, *41*(1), 53–61.
- Saper, C. B., Scammell, T. E., & Lu, J. (2005). Hypothalamic regulation of sleep and circadian rhythms. *Nature*, *437*(7063), 1257–1263. <https://doi.org/10.1038/nature04284>
- Sassenhagen, J., & Draschkow, D. (2019). Cluster-based permutation tests of MEG/EEG data do not establish significance of effect latency or location. *Psychophysiology*, *56*(6), e13335. <https://doi.org/10.1111/psyp.13335>
- Shapiro, L. A. (2010). *Embodied cognition. New problems of philosophy*. Routledge. <http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10416730>
- Steyer, R., Notz, P., Schwenkmezger, P., & Eid, M. (1997). *Der Mehrdimensionale Befindlichkeitsfragebogen* (1st ed.). Hogrefe.
- Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, *54*(5), 768–777. <https://doi.org/10.1037/0022-3514.54.5.768>
- Stutz, J., Eiholzer, R., & Spengler, C. M. (2019). Effects of evening exercise on sleep in healthy participants: A systematic review and meta-analysis. *Sports Medicine*, *49*(2), 269–287. <https://doi.org/10.1007/s40279-018-1015-0>
- Tarvainen, M. P., Niskanen, J.-P., Lipponen, J. A., Ranta-Aho, P. O., & Karjalainen, P. A. (2014). Kubios HRV - Heart rate variability analysis software. *Computer Methods and Programs in Biomedicine*, *113*(1), 210–220. <https://doi.org/10.1016/j.cmpb.2013.07.024>
- Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., Fazio, F., Rizzolatti, G., Cappa, S. F., & Perani, D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, *17*(2), 273–281. <https://doi.org/10.1162/0898929053124965>
- Thibault, R. T., & Raz, A. (2016). Imaging posture veils neural signals. *Frontiers in Human Neuroscience*, *10*, 520. <https://doi.org/10.3389/fnhum.2016.00520>
- Tukey, J. W. (1977). *Exploratory data analysis. Addison-Wesley series in behavioral science Quantitative methods*. Addison-Wesley.
- Wagenmakers, E.-J., Beek, T., Dijkhoff, L., Gronau, Q. F., Acosta, A., Adams, R. B., Albohn, D. N., Allard, E. S., Benning, S. D., Blouin-Hudon, E.-M., Bulnes, L. C., Caldwell, T. L., Calin-Jageman, R. J., Capaldi, C. A., Carfagno, N. S., Chasten, K. T., Cleeremans, A., Connell, L., DeCicco, J. M., ... Zwaan, R. A. (2016). Registered replication report: Strack, Martin, & Stepper (1988). *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *11*(6), 917–928. <https://doi.org/10.1177/1745691616674458>
- Widmann, A., Schröger, E., & Maess, B. (2015). Digital filter design for electrophysiological data—A practical approach. *Journal of Neuroscience Methods*, *250*, 34–46. <https://doi.org/10.1016/j.jneumeth.2014.08.002>

- World Medical Association. (2001). World medical association declaration of Helsinki. Ethical principles for medical research involving human subjects. *Bulletin of the World Health Organization*, 79(4), 373–374.
- Wuyts, J., De Valck, E., Vandekerckhove, M., Pattyn, N., Bulckaert, A., Berckmans, D., Haex, B., Verbraecken, J., & Cluydts, R. (2012). The influence of pre-sleep cognitive arousal on sleep onset processes. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, 83(1), 8–15. <https://doi.org/10.1016/j.ijpsycho.2011.09.016>
- Yang, J., & Shu, H. (2016). Involvement of the motor system in comprehension of non-literal action language: A meta-analysis study. *Brain Topography*, 29(1), 94–107. <https://doi.org/10.1007/s10548-015-0427-5>
- Zwaan, R. A., van der Stoep, N., Guadalupe, T., & Bouwmeester, S. (2012). Language comprehension in the balance: The robustness

of the action-compatibility effect (ACE). *PLoS One*, 7(2), e31204. <https://doi.org/10.1371/journal.pone.0031204>

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

How to cite this article: Hülsemann MJ, Rasch B. Embodiment of sleep-related words: Evidence from event-related potentials. *Psychophysiology*. 2021;58:e13824. <https://doi.org/10.1111/psyp.13824>