



Phonetic Skills and Verbal Memory Capacity Predict Phonetic-based Word Learning: An Event-related Potential Study

Stefan Elmer^{1,2}, Eva Dittinger³ , Julia Brocchetto¹, Clément François⁴, Mireille Besson³, Lutz Jäncke^{1,5*}, and Antoni Rodríguez-Fornells^{2,6,7*}

Abstract

■ The learning of new words is a challenge that accompanies human beings throughout the entire life span. Although the main electrophysiological markers of word learning have already been described, little is known about the performance-dependent neural machinery underlying this exceptional human faculty. Furthermore, it is currently unknown how word learning abilities are related to verbal memory capacity, auditory attention functions, phonetic discrimination skills, and musicality. Accordingly, we used EEG and examined 40 individuals, who were assigned to two groups (low [LPs] and high performers [HPs]) based on a median split of word learning performance, while they completed a phonetic-based word learning task. Furthermore, we collected behavioral data during an attentive listening and a phonetic discrimination task with the same stimuli to address relationships between auditory attention

and phonetic discrimination skills, word learning performance, and musicality. The phonetic-based word learning task, which also included a nonlearning control condition, was sensitive enough to segregate learning-specific and unspecific N200/N400 manifestations along the anterior–posterior topographical axis. Notably, HPs exhibited enhanced verbal memory capacity and we also revealed a performance-dependent spatial N400 pattern, with maximal amplitudes at posterior electrodes in HPs and central maxima in LPs. Furthermore, phonetic-based word learning performance correlated with verbal memory capacity and phonetic discrimination skills, whereas the latter was related to musicality. This experimental approach clearly highlights the multifaceted dimensions of phonetic-based word learning and is helpful to disentangle learning-specific and unspecific ERPs. ■

INTRODUCTION

Learning the meaning of new words is an intriguing task that continues to attract interest from a broad spectrum of disciplines, including education, linguistics, psychology, and neuroscience (Bahramlou & Esmaeili, 2019; Eiteljoerge, Adam, Elsner, & Mani, 2019; Rodríguez-Fornells, Cunillera, Mestres-Misse, & de Diego-Balaguer, 2009). The reason for such widespread interest is possibly grounded in the multifaceted nature of the perceptual and cognitive functions involved. In fact, to rapidly acquire new words, learners need to engage several resources, including phonetic processing and discrimination, attention, as well as verbal and associative memory functions (Takashima, Bakker, van Hell, Janzen, & McQueen, 2017; de Diego-Balaguer,

Martinez-Alvarez, & Pons, 2016; Dittinger et al., 2016). Notwithstanding the fast proliferation of research on word learning, we currently know relatively little about the contribution of these processing resources and the corresponding neural correlates to the word acquisition process (Rodríguez-Fornells et al., 2009). Hence, one of the biggest challenges that face research on word learning is the identification of reliable neural markers as well as psychological, perceptual, and biographical key variables that modulate word learning.

The EEG technique is particularly suited to grasp the fast neural dynamics associated with learning in different domains (Rueda-Delgado, Heise, Daffertshofer, Mantini, & Swinnen, 2019; Nan et al., 2018; Langer, von Bastian, Wirz, Oberauer, & Jancke, 2013). It is noteworthy to mention that most of the previous EEG studies on word learning focused on event-related potentials (ERPs) with the attempt to objectify the neural manifestations underlying the acquisition and consolidation process (François, Cunillera, Garcia, Laine, & Rodríguez-Fornells, 2017; Bakker, Takashima, van Hell, Janzen, & McQueen, 2015; Borovsky, Kutas, & Elman, 2010; Dobel, Lagemann, & Zwitserlood, 2009). In this context, the N200 and N400 ERP components have commonly been identified as suitable markers for capturing

¹University of Zurich, Switzerland, ²Bellvitge Biomedical Research Institute, Barcelona, Spain, ³Université Publique de France, CNRS & Aix-Marseille University, ⁴Aix-Marseille University, Aix-en-Provence, France, ⁵University Research Priority Program (URRP), Zurich, Switzerland, ⁶University of Barcelona, Spain, ⁷Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain

*Shared last authorship.

the reconfiguration of hierarchical neural networks involved in word-to-meaning mapping (Rasamimanana, Barbaroux, Cole, & Besson, 2020; Francois et al., 2017; Dittinger et al., 2016; Junge, Cutler, & Hagoort, 2012). Based on current knowledge, the auditory N200 component is thought to mirror phonetic encoding and categorization processes (Friedrich & Friederici, 2008; Connolly & Phillips, 1994), whereas the N400 response has primarily been associated with the building up of lexical–semantic representations (Kutas & Federmeier, 2011; Borovsky et al., 2010). However, the intrinsic meaning of these two ERPs in word learning tasks is still a matter of debate (Dittinger et al., 2016). In fact, although the N400 component has been shown to be sensitive to the word–meaning acquisition process, its exclusive sensitivity to lexical–semantic information is controversial (Kim, Oines, & Miyake, 2018; Kutas & Federmeier, 2011; Lau, Phillips, & Poeppel, 2008). In particular, the N400 waveform has previously been associated with short-term and working memory functions (Hagoort, 2014; Kutas & Federmeier, 2011), associative memory (Hagoort, 2014), and the storage of word representations in episodic memory (Rodriguez-Fornells et al., 2009; Wagner et al., 1998), and has also been shown to be modulated by attention (Getzmann, Jasny, & Falkenstein, 2017; Wang, Zhang, & Liu, 2017; Garagnani, Wennekers, & Pulvermuller, 2008). The processes mirrored by the N200 component are also somehow elusive in that this waveform can, for example, be influenced by attention (Patel & Azzam, 2005; Schroger, 1993) and conflict monitoring processes (Enriquez-Geppert, Konrad, Pantev, & Huster, 2010; Dimoska, Johnstone, & Barry, 2006). Furthermore, the N200 has been shown to reflect early lexical selection and initial word–form recognition (van den Brink, Brown, & Hagoort, 2001). Consequently, the functional role of these two ERP components in word learning tasks can only be properly inferred from experimental designs that enable to discriminate between learning-specific and unspecific ERP modulations related to cognitive load, attention, short-term memory, or working memory load.

Up to now, associative (Dittinger, Scherer, Jancke, Besson, & Elmer, 2019; Francois et al., 2017; Dobel et al., 2009; Friedrich & Friederici, 2008; von Koss Torkildsen et al., 2008) and contextual learning tasks (Batterink & Neville, 2011; Borovsky et al., 2010; Frishkoff, Perfetti, & Collins-Thompson, 2010; Mestres-Misse, Rodriguez-Fornells, & Munte, 2007) have most commonly been used to objectify N200 and N400 dynamics underlying the word–meaning acquisition process. Associative learning refers to the ability to acquire the meaning of new items by coupling them with figural representations (word-to-figure mapping) or words (word-to-word mapping). In contrast, contextual learning implies that the linguistic context in which a new word appears can be used to infer its meaning. For example, Mestres-Missé et al. (Mestres-Misse et al., 2007) applied a word learning task where the participants were required to discover the meaning of novel words embedded in

three consecutive hierarchically organized sentences. Thereby, the authors compared neural activity between the newly learned words, words that could not be deduced from the context, and existing words. After only three exposures, N400 responses to novel words could no longer be distinguished from those induced by real words over central and posterior scalp sites, whereas such an effect was not apparent for words that could not be learned based on contextual information.

To date, several authors have studied the neural bases of the word–meaning acquisition process using associative learning tasks (Bakker-Marshall et al., 2018; Friedrich & Friederici, 2008; von Koss Torkildsen et al., 2008). For example, in a series of experiments with children (Dittinger, Chobert, Ziegler, & Besson, 2017), young adults (Dittinger et al., 2016), students suffering from dyslexia (Rasamimanana et al., 2020), and older adults (Dittinger et al., 2019), Besson et al. presented picture–word associations (words = Thai monosyllabic words) to the participants and evaluated N200 and N400 reconfigurations during the word–meaning acquisition process. Furthermore, in a subsequent matching task, the authors tested learning performance and evaluated EEG traces while participants had to distinguish between match and mismatch trials based on the previously learned associations. Importantly, because several reports postulated a positive influence of music training on speech perception (Zendel, West, Belleville, & Peretz, 2019; Elmer, Meyer, & Jancke, 2012; Besson, Chobert, & Marie, 2011) and cognitive functions (Strait, Slater, O’Connell, & Kraus, 2015; Zuk, Benjamin, Kenyon, & Gaab, 2014; Moreno et al., 2011), the authors also evaluated behavioral and electrophysiological differences between musically trained and untrained individuals. In this context, it is important to emphasize that using Thai words as stimulus material for learning the meaning of new words also implies that the participants have to learn to distinguish the new phonetic contrasts. Accordingly, the ability to discriminate the new Thai phonemes constitutes a fundamental prerequisite for word learning.

In general, these experiments revealed highly consistent spatio-temporal N200 and N400 patterns across all examined cohorts. In fact, (1) in the learning phase, N200 and N400 responses were larger over frontal and central compared to posterior electrodes, and (2) both ERP components dynamically increased across the learning blocks. Furthermore, (3) in the “matching task,” N200 and N400 responses spatially shifted from anterior-central to posterior electrodes. Taking into account the effect of music training, (4) musicians and musically trained children showed a larger increase in N400 amplitudes in the second compared to the first block of the learning phase. In the matching task, only musically trained individuals exhibited central-posterior N200 and N400 effects manifested by larger amplitudes in response to mismatch compared to match trials. Finally, (5) adult musicians also performed better than untrained individuals and demonstrated increased N100 amplitudes in a phonetic categorization task

that consisted of the same Thai monosyllabic words and was administered before the word learning phase.

These pivotal studies examining word learning through picture–word associations paved the way for a better understanding of the electrophysiological dynamics underlying word–meaning acquisition. Furthermore, these studies were innovative in that they described relationships between musical practice, word learning, and phonetic processing. However, some important research questions have not yet been systematically addressed or remain unresolved. In particular, (1) it is unclear whether the observed anterior-central N200 and N400 distributions in the learning phase specifically reflected the word–meaning acquisition process, or whether these two ERPs were possibly comodulated by the engagement of task-related cognitive functions like attention, short-term memory, and working memory demands. Furthermore, because Dittinger et al. (Dittinger et al., 2016, 2017, 2019) did not test performance immediately after the learning phase, (2) it is obscure whether the observed shift of N200 and N400 components from anterior-central to posterior electrodes in the “matching task” mirrored incremental learning, more efficient integration of the learned words into lexical–semantic memory, or simply active retrieval processes. Moreover, although Dittinger and coworkers provided convergent evidence for an influence of music training on word learning (Dittinger et al., 2016, 2017, 2019), (3) it is still unresolved whether these effects were driven by professional music training per se, musicality, or facilitated learning. Finally, (4) the question of whether there is a possible relationship between phonetic-based word learning performance, auditory attention, verbal memory functions, and phonetic discrimination ability has not yet been addressed.

In order to shed light on these open questions, we integrated the Thai words used by Dittinger et al. (Dittinger et al., 2016) into a new experimental design composed of an attentive listening task, a phonetic discrimination task, and an optimized phonetic-based word learning paradigm. The attentive listening task consisted of detecting infrequent Thai word repetitions and had the purpose of testing auditory encoding, attention, and phonetic discrimination functions (Fritz, Elhilali, David, & Shamma, 2007; Michie, Bearpark, Crawford, & Glue, 1990), whereas the phonetic discrimination task served to examine the ability of the participants to distinguish the words later used in the phonetic-based word learning task (Elmer, Greber, Pushparaj, Kuhn, & Jancke, 2017). The behavioral data of the attentive listening and discrimination tasks were used to uncover relationships between auditory attention and phonetic discrimination skills, word learning performance, and musicality. Most importantly, the main innovation of the phonetic-based word learning task was that we also included a control condition consisting of words that were randomly associated with different pictures, so that participants could not learn their meaning (nonlearning condition). This allowed us to

tease apart learning-specific (related to the word–meaning acquisition) and unspecific (reflecting task-related cognitive load) N200/N400 manifestations. Finally, performance was directly tested after the learning phase.

Drawing on this background, we treated musicality (tonal and rhythmic scores) as a continuous variable and recruited 40 participants who were assigned to two groups based on a median split of the word learning performance (high performers [HPs] and low performers [LPs]). Such an approach offers the advantage of distinguishing between the general neural principles underlying phonetic-based word learning and performance-dependent ERP manifestations. Furthermore, it enables to verify whether word learning performance is related to auditory attention and discrimination skills and whether musicality and verbal memory capacity possibly act as moderator variables.

METHODS

Participants

Forty participants in the age range of 20–40 years (mean age = 25.46 years, $SD = 4.9$ years) and without neurological or psychological deficits were recruited for the study. Because of extensive EEG artifacts, one participant had to be excluded from the analyses. All participants were consistently right-handed (Annett, 1970), native German speakers, and none of them were simultaneous or early bilingual. We deliberately did not exclude individuals with current or past musical practice and treated musicality as a continuous variable for correlation analyses. Furthermore, to examine performance-dependent behavioral and electrophysiological data, the participants were assigned to two groups based on a median split of word learning performance according to the percentage of correctly identified learned and nonlearned items. The participants were paid for participation and gave informed written consent in accordance with the procedures of the local ethics committee and the declaration of Helsinki.

Pure-Tone Audiometry, Musicality, and History of Musical Practice

Pure-tone audiometry (MAICO Diagnostic GmBh) was tested in the frequency range of 250–8000 Hz. The participants demonstrated an unremarkable audiological status in that all tested frequencies could be heard below a threshold of 30 dB. Musicality was assessed by the Advanced Measure of Music Audition (AMMA) test (Gordon, 1989) consisting of comparing pairs of piano melodies and deciding whether the melodies were equivalent, rhythmically different, or tonally different. The total AMMA score (i.e., both rhythmic and tonal subtests) was used as a continuous variable for correlative analyses. Furthermore, biographical data on the history of musical

practice were collected using an in-house questionnaire (Elmer et al., 2012).

Cognitive Capabilities

To rule out between-group differences in general intelligence, all participants completed a standardized form of the KAI test (*Kurztest für Allgemeine Basisgrößen der Informationsverarbeitung*; Lehrl, Blaha, & Fischer, 1992). During the KAI test, the participants had to read aloud meaningless sequences of 20 letters as quickly as possible and to repeat auditory-presented letters and digits increasing in length (up to nine items). The KAI test has been shown to correlate about $r = .7$ with global IQ in healthy adults (Lehrl et al., 1992). Moreover, to examine whether word-acquisition capacity was possibly related to verbal memory functions, each participant also completed the VLMT test (*verbaler Lern- und Merkfähigkeitstest*; Helmstaedter & Lux, 2001). This procedure is used as a reliable measure for assessing short-term verbal memory functions and consisted of remembering as many auditory-presented words as possible from a list of 15 items that was read out once by the investigator.

Auditory Stimuli

The Thai monosyllabic consonant–vowel words used in the phonetic-based word learning task were taken from a corpus of 12 units (/ba0/, /ba:0/, ba:1, /ba1/, /pa0/, /pa:0/, /pa:1/, /pa1/, /p^ha0/, /p^ha1/, /p^ha:0/, and /p^ha:1/) previously created by Dittinger et al. (Dittinger et al., 2016). In order to reproduce natural speech variability, four versions of each word were recorded by a native female Thai speaker. Sound pressure level was normalized across all items to a mean level of 70 dB (according to a 0-dB reference) using the Praat software (www.fon.hum.uva.nl/praat/).

The auditory stimuli consisted of natural Thai monosyllabic words with short (/ba1/, /pa1/, and /p^ha1/; 261 msec on average) and long vowel duration (/ba:1/, /pa:1/, /p^ha:1/, /ba:0/, /pa:0/, and /p^ha:0/; 531 msec on average), with low-tone (/ba1/, /pa1/, /p^ha1/, /ba:1/, /pa:1/, and /p^ha:1/; F₀ = 175 Hz on average) and mid-tone vowels (/ba:0/, /pa:0/, and /p^ha:0/; F₀ = 218 Hz on average) as well as with vowels varying in voicing (/ba1/, /ba:1/, and /ba:0/, VOT = –144 msec vs. /pa1/, /pa:1/ and /pa:0/, VOT = 3 msec) and in aspiration contrasts (/pa1/, /pa:1/, and /pa:0/, VOT = 3 msec vs. /p^ha1/, /p^ha:1/, and /p^ha:0/, VOT = 77 msec).

Based on pilot experiments, this corpus of 12 words was reduced to 10 to optimize the word learning curve. Accordingly, because aspirated syllables exist in the German language, only two out of four of these stimuli were presented to each participant. However, to guarantee a certain degree of variability, we assigned the four aspirated words to two different pools of stimuli that consisted of the same eight words without aspiration but differed in two aspirated items, namely, /P^ha1/ and /P^ha0/

versus /P^ha:1/ and /P^ha:0/. These two pools of 10 stimuli were pseudorandomly counterbalanced across participants and used in the attentive listening, phonetic discrimination, and word learning tasks.

Visual Stimuli

For each of the 12 words used in the word learning experiment, we selected 10 similar variations of black and white pictures (e.g., 10 pictures of dogs). These pictures represented concrete living and nonliving objects with a high prototypicality for the following semantic categories: fruits (apple), animals (dog), weapons (pistol), office supplies (pencil), body parts (arm), clothes (trousers), vehicles (car), tools (hammer), buildings (house), kitchen equipment (fork), musical instruments (tambourine), and furniture (table). These different pictures were chosen from the Internet based on previous studies that evaluated objects' prototypicality of different semantic categories (Barbarotto, Laiacona, Macchi, & Capitani, 2002; Maess, Friederici, Damian, Meyer, & Levelt, 2002). All pictures were matched in size (width = 16 cm; height = 20 cm) and presented in the middle of a computer screen.

Experimental Procedure

During the experimental session, the participants were seated within a Faraday cage in a comfortable chair at about one meter from a computer screen. Auditory stimuli were presented through HiFi headphones (HD590, Sennheiser Electronic GmbH) at about 70-dB sound pressure level. Visual and auditory stimuli presentation, as well as the collection of behavioral responses, was controlled by the Presentation software (Version 11.0, Neurobehavioral Systems).

Experimental Conditions

Attentive Listening Task

The two pools of 10 auditory stimuli were pseudorandomly presented across participants. For both pools, each of the 10 words was presented 13 times, resulting in a total of 130 stimuli. Furthermore, each of the 10 words was presented twice in a row and the participants had to identify word repetitions by pressing a response button (10 repetitions out of 130 trials). Because the different stimuli slightly differed in duration, SOA was not jittered but kept constant at 1500 msec. The entire task had a duration of 3.25 min and served to assess the behavioral correlates of auditory attention functions and word encoding (Fritz et al., 2007; Michie et al., 1990).

Phonetic Discrimination Task

In the phonetic discrimination task, we used the same two pools of 10 words presented in the attentive listening

task. In each trial, participants had to decide whether the word pairs were same or different by pressing one out of two response buttons. Each participant was exposed to 160 word pairs (80 same and 80 different) presented with an SOA of 2000 msec and a trial duration of 3500 msec. Each of the 10 words was presented 16 times, 8 in the same and 8 in the different condition. Furthermore, 25% of the different trials varied in duration, 25% in aspiration, 25% in pitch, and 25% in voicing. In particular, for trials containing changes in pitch, duration, or VOT, we used minimal pairs. Otherwise, because only the phoneme /Pha/ contained an aspiration cue, trials with changes in aspiration could also vary on another acoustic dimension. The total task duration was 9.2 min.

Phonetic-based Word Learning Task

The phonetic-based word learning task consisted of two successive blocks of 4.15 min each, and the participants were explicitly instructed to learn the meaning of the new words based on picture–word associations. In each block, half of the words of each pool (five) were consistently associated (learning condition) with variations of the same pictures (e.g., different pictures of a fork), whereas the other half was randomly coupled with different visual items (non-learning condition). Participants were not aware of the non-learning condition. Furthermore, to exclude the influence of stimulus material on word learning, the words used in the learning and nonlearning conditions were counterbalanced across participants. In particular, the words that in one version were consistently associated with the same pictures were presented with inconsistent ones across participants in the other version. Accordingly, for each of the two pools of stimuli, we created two different versions (Pool 1 Version 1: /ba1/, /ba:0/, /pa1/, /pa:0/, /p^ha:1/; Pool 1 Version 2: /ba:1/, /ba0/, /pa:1/, /pa0/, /p^ha:0/; Pool 2 Version 1: /ba1/, /ba:0/, /pa1/, /pa:0/, /p^ha0/; Pool 2 Version 2: /ba:1/, /ba0/, /pa:1/, /pa0/, /p^ha1/).

In the phonetic-based word learning task, participants were exposed to 1 of the 10 pictures that were presented for 2000 msec and followed, 800 msec after picture presentation onset (SOA), by one of the words (trial duration = 2500 msec). Each of the two blocks consisted of 100 trials, and every single word of the learning (five words) and nonlearning (five words) condition was presented 10 times in association with 10 variations of the same pictures (learning condition) or two variations of each of the five inconsistent pictures (nonlearning condition).

Immediately after the learning phase, performance was tested using a forced-choice (FC) task. In the FC task, four pictures (two of the learning and two of the non-learning condition) were simultaneously presented side by side on the screen for 6500 msec (trial duration). Words were presented 800 msec after the onset of the pictures, and the participants had to select the picture that matched the meaning of the learned word by pressing the corresponding response button. Furthermore,

the participants were instructed to press an additional response button when they thought that they had not learned the meaning of the presented word. This additional response key served as correct response for the words of the nonlearning condition. The FC task consisted of 30 trials, each of the five words of the learning and nonlearning condition was presented 3 times, and each picture (including its variations) was presented 12 times. The test phase had a duration of 3.25 min.

EEG Data Acquisition and Preprocessing

The EEG was recorded at a sampling rate of 1000 Hz with a high-pass filter of 0.1 Hz using an EEG amplifier (Brain Products, GmbH). Thirty-two active Ag/Cl electrodes were located at standard positions according to the international 10–20 system, the reference electrode was placed on the tip of the nose, and electrode impedance was kept below 10 k Ω . The EEG data were preprocessed using the Brain Vision Analyzer software (Version 2.1.0). In particular, the data of the phonetic-based word learning task were rereferenced off-line to the averaged left and right mastoids and filtered with a bandpass filter of 0.1–30 Hz (slope of 48 dB/oct) and a Notch filter of 50 Hz. Furthermore, an independent component analysis was used to identify and correct vertical and horizontal ocular movements, and the remaining artifacts were automatically removed according to a maximum–minimum criterion of 100 μ V. For each participant, single epochs time-locked to the onset of the words were extracted (from –200 to 600 msec), averaged and baseline-corrected (from –200 to 0 msec). Individual averages were then used to compute grand averages for the HP and LP groups as well as for the whole sample of participants.

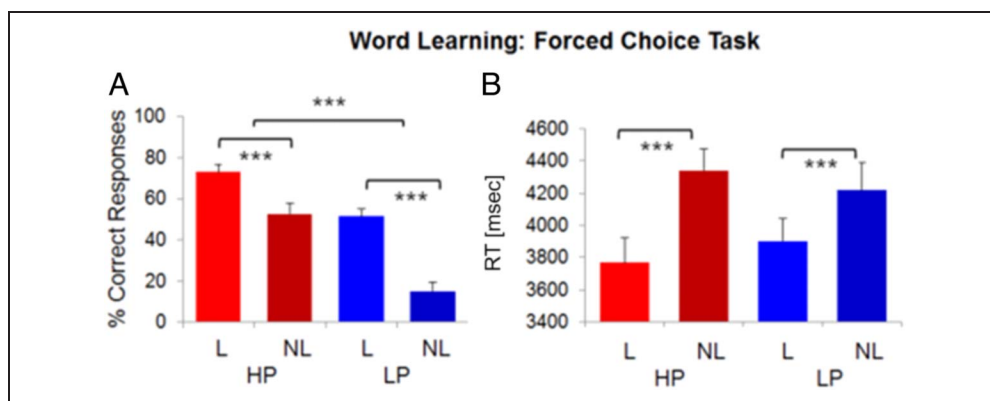
EEG Analyses

In the phonetic-based word learning task, we evaluated mean amplitudes of the N200 (250–350 msec) and N400 (350–550 msec) components in three ROIs located at anterior (F3, Fz, F4), central (C3, Cz, C4), and posterior (P3, Pz, P4) scalp sites. This procedure was chosen according to previous work showing a shift on N200 and N400 responses from anterior to posterior scalp sites as a function of word learning (Ditinger et al., 2016, 2017, 2019). Furthermore, to disentangle learning-specific and unspecific N200 and N400 effects, we computed difference values of the learning and nonlearning conditions in the time window of the N200 (250–350 msec) and N400 (350–550 msec) components.

Statistical Analyses

All analyses were performed using parametric statistics implemented in the IBM SPSS Statistics 22 software. The evaluation of the biographical, psychometric, and behavioral data was carried out using *t* tests for independent

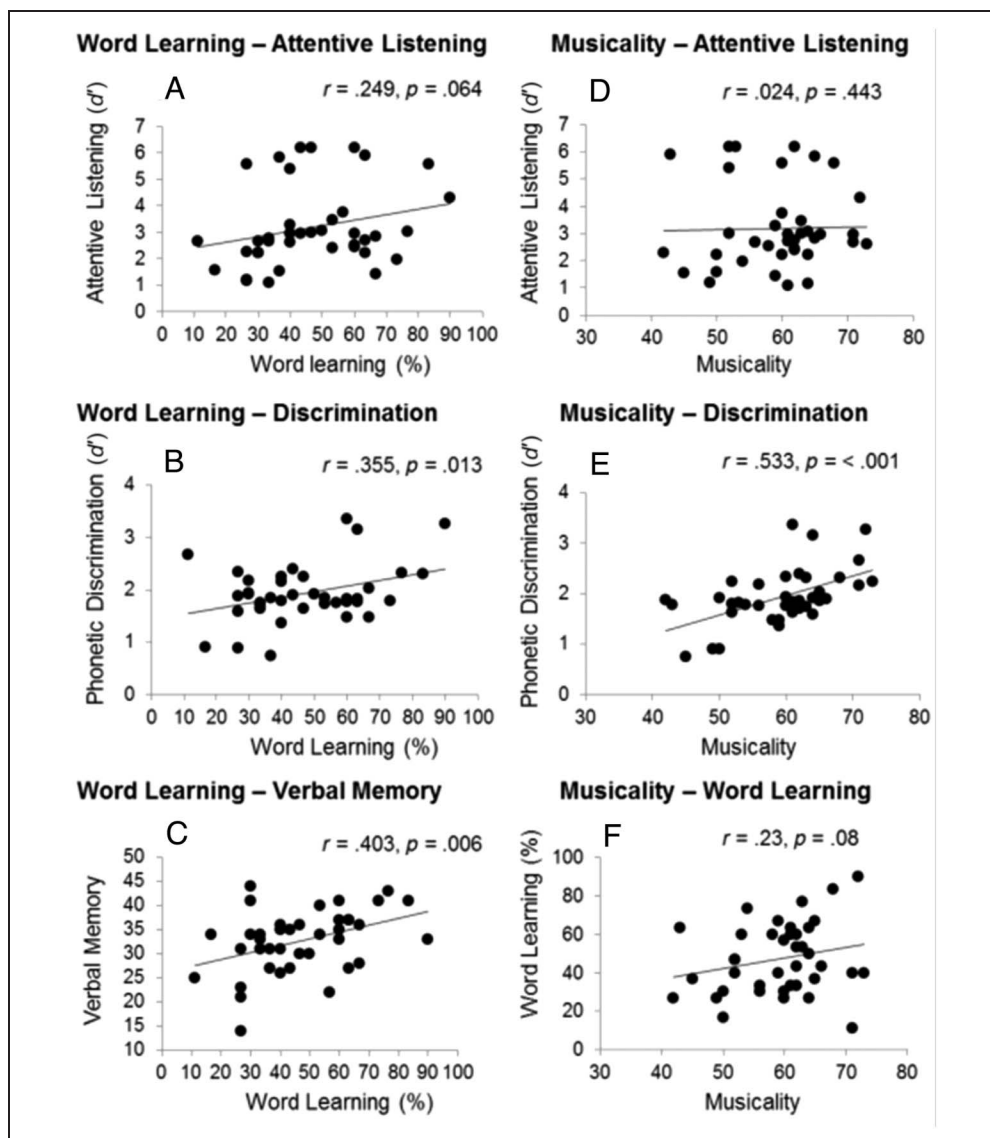
Figure 1. Percentage of correct responses (A) and RTs (B) in the phonetic-based word learning task are shown separately for the HP and LP groups. The bars indicate the standard error of mean. $** p < .01$, $*** p < .001$. L = learning condition; NL = nonlearning condition.



samples (two-tailed) or ANOVAs (repeated measurements). Furthermore, for the attentive listening and phonetic discrimination tasks, we computed d' values that were used for correlation analyses (Stanislaw & Todorov, 1999).

In the phonetic-based word learning task, mean N200/N400 amplitudes were examined using separate ANOVAs (repeated measurements) with the between-subjects factor of Group and the within-subject factors of Block (first and second), Condition (learning and

Figure 2. Behavior–behavior correlations between phonetic-based word learning performance (percentage of correct responses) and attentive listening (A, d' values), phonetic-based word learning performance (percentage of correct responses) and phonetic discrimination (B, d' values), and phonetic-based word learning performance (percentage of correct responses) and verbal memory (C). The right column depicts correlations between musicality and attentive listening (D, d' values), musicality and phonetic discrimination (E, d' values), and musicality and phonetic-based word learning (percentage of correct responses, F).



nonlearning), and ROI (anterior, central, and posterior). Furthermore, ANOVAs (two groups, two blocks, and three ROI) were also used to evaluate the N200 and N400-related difference values generated by subtracting the nonlearning from the learning condition. The purpose of this additional analysis was to correct the data for learning-unspecific task-related effects. Significant main effects and interactions were further inspected using post hoc *t* tests or ANOVAs.

Correlative analyses in the whole sample of participants were used to carve out behavioral relationships between word learning performance and auditory attention functions (d' values), verbal memory capacity (VLMT), and phonetic discrimination skills (d' values). Moreover, we correlated the musicality score (AMMA test) with auditory attention functions (d' values), phonetic discrimination skills (d' values), and word learning performance. Based on the fact that in this context the examination of negative correlations does not make sense, correlations were computed in a one-tailed manner using Pearson's r (corrected for multiple comparisons).

RESULTS

Biographical Data, Musical Aptitudes, and Cognitive Capabilities

t Tests for independent samples revealed that the two groups (HPs and LPs) did not differ in age, $t(37) = -0.144, p = .887$; IQ (KAI, $t(37) = 0.57, p = .572$); musicality (total AMMA score, $t(37) = 0.438, p = .664$); or cumulative number of hours of music training, $t(37) = -0.44, p = .663$. However, as expected, statistical analysis of the VLMT test showed that HPs were characterized by a higher verbal memory capacity compared to LPs, $t(37) = 2.063, p = .046$, Cohen's $d = 0.66$.

Behavioral Data

Phonetic-based Word Learning: Forced Choice Task

The behavioral data of the FC task that followed the word learning task (percentage of correct responses and RT; Figure 1) were analyzed using 2×2 ANOVAs (two groups and two conditions). The evaluation of percentage of

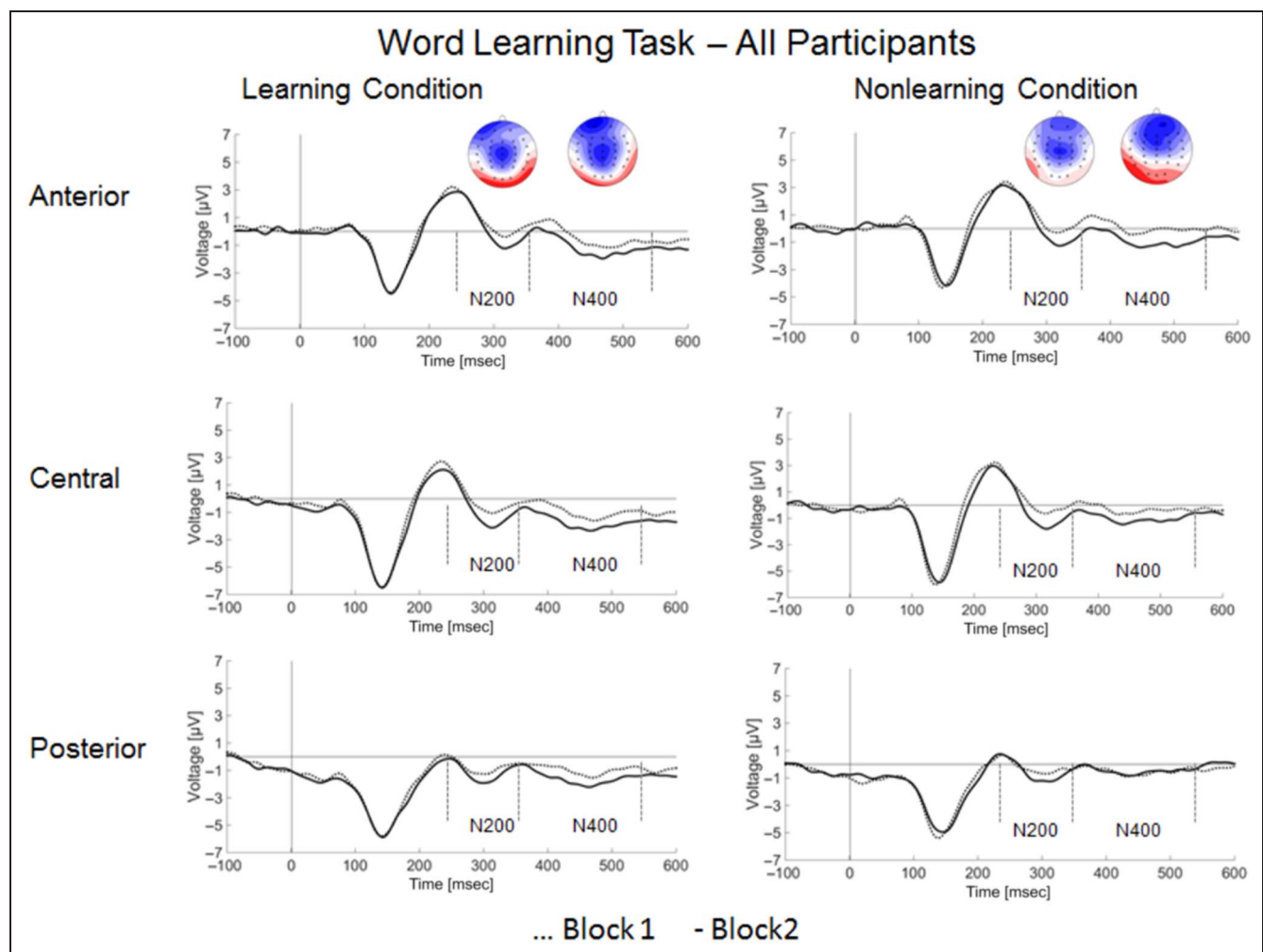


Figure 3. The EEG traces in the whole sample of participants are shown separately for the learning (left side) and nonlearning (right side) conditions at anterior (first row), central (second row), and posterior (third row) electrodes. Dotted line = first block, solid line = second block. The topographical maps depict voltage differences between the two blocks in the N200 and N400 time windows.

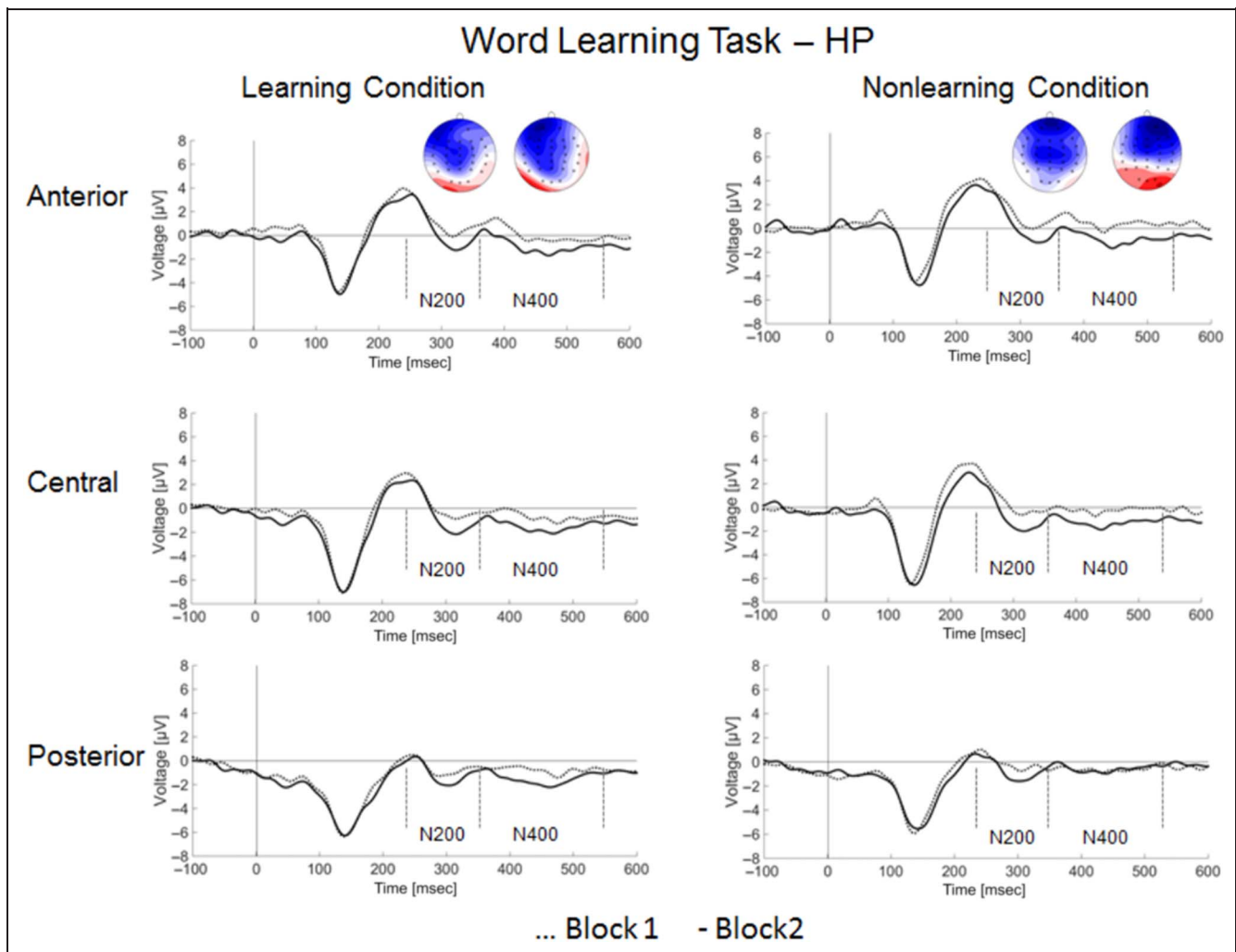


Figure 4. The EEG traces of the HP group are shown separately for the learning (left side) and nonlearning (right side) conditions at anterior (first row), central (second row), and posterior (third row) electrodes. Dotted line = first block, solid line = second block. The topographical maps depict voltage differences between the two blocks in the N200 and N400 time windows.

correct responses yielded main effects of Condition, $F(1, 37) = 29.42, p < .001$; partial $\eta^2 = .443$, and Group, $F(1, 37) = 95.185, p < .001$; partial $\eta^2 = .72$, whereas the analysis of RT showed a main effect of Condition, $F(1, 37) = 34.336, p < .001$; partial $\eta^2 = .482$. HPs were characterized by overall more correct responses compared to LPs. However, because the two groups were defined according to a median split of the word learning performance, this effect was expected. Furthermore, all participants demonstrated more correct responses and shorter RTs in the learning than in the nonlearning condition.

Behavior–Behavior Relationships

Within the whole sample of participants, we correlated phonetic-based word learning performance in the FC task (one-tailed, corrected p value for three correlations = .016) with verbal memory capacity (VLMT), d' values of the attentive listening task, and d' values of the phonetic

discrimination task. Results (Figure 2) revealed that phonetic-based word learning performance in the FC task was positively related to phonetic discrimination skills ($r = .355, p = .013$) and verbal memory capacity ($r = .403, p = .006$).

As an additional goal of the study, we tested possible relationships (one-tailed, corrected p value for three correlations = .016) between musicality (total AMMA score), d' values of the attentive listening task, d' values of the phonetic discrimination task, and phonetic-based word learning performance in the FC task. The correlative analyses (Figure 2) only revealed that musicality was positively correlated with phonetic discrimination abilities ($r = .533, p < .001$).

Electrophysiological Data

Phonetic-based Word Learning Task

In the phonetic-based word learning task, mean N200 and N400 amplitudes were examined with separate $2 \times 2 \times 2 \times$

3 ANOVAs (two groups [HPs vs. LPs], two conditions [learning vs. nonlearning], two blocks [B1 vs. B2], and three ROIs [anterior vs. central vs. Posterior]). The ANOVA performed on N200 amplitudes revealed main effects of Block, $F(1, 37) = 9.874, p = .003$; partial $\eta^2 = .211$, and ROI, $F(1, 37) = 4.194, p = .038$; partial $\eta^2 = .102$, as well as significant Condition \times ROI, $F(1, 37) = 4.134, p = .037$; partial $\eta^2 = .101$, and Block \times ROI, $F(1, 37) = 3.784, p = .05$; partial $\eta^2 = .093$, interaction effects. The main effect of Block was because of increased N200 amplitudes in Block 2 compared to Block 1 (Figure 3), whereas t tests for dependent samples used to disentangle the main effect of ROI revealed larger N200 amplitudes at the central compared to the anterior ROI (anterior-central: $t(38) = 3.783, p = .001$; anterior-posterior: $t(38) = 1.89, p = .066$; central-posterior: $t(38) = -0.014, p = .989$, Figure 3). Furthermore, post hoc t tests targeting at comparing the two conditions at anterior, central, and posterior ROIs (Condition \times ROI interaction) reached significance at posterior electrodes only (anterior: $t(38) = -0.345, p = .732$; central: $t(38) = -1.119, p = .062$; posterior: $t(38) =$

$-2.252, p = .03$). This effect was related to larger N200 amplitudes in the learning compared to the nonlearning condition (Figure 3). Finally, post hoc t tests used to disentangle the Block \times ROI interaction revealed that, at anterior and central electrodes, the N200 amplitude was larger in Block 2 compared to Block 1 (anterior: $t(38) = 3.102, p = .004$; central: $t(38) = 3.783, p = .001$; posterior: $t(38) = 1.839, p = .074$; Figure 3).

Statistical analysis of N400 amplitudes yielded main effects of Condition, $F(1, 37) = 9.454, p = .004$; partial $\eta^2 = .204$, and Block, $F(1, 37) = 5.165, p = .029$; partial $\eta^2 = .122$, as well as Block \times ROI, $F(1, 37) = 4.051, p = .044$; partial $\eta^2 = .099$; Condition \times Block \times ROI, $F(1, 37) = 3.746, p = .017$; partial $\eta^2 = .127$; and Group \times Condition \times Block \times ROI interactions (quadratic interaction: $F(1, 37) = 5.098, p = .03$; partial $\eta^2 = .121$). The main effect of Condition originated from larger N400 amplitudes in the learning compared to the nonlearning condition (Figure 3), whereas the main effect of Block was related to increased N400 responses in Block 2 compared to Block 1 (Figure 3). Post hoc t tests used to examine the

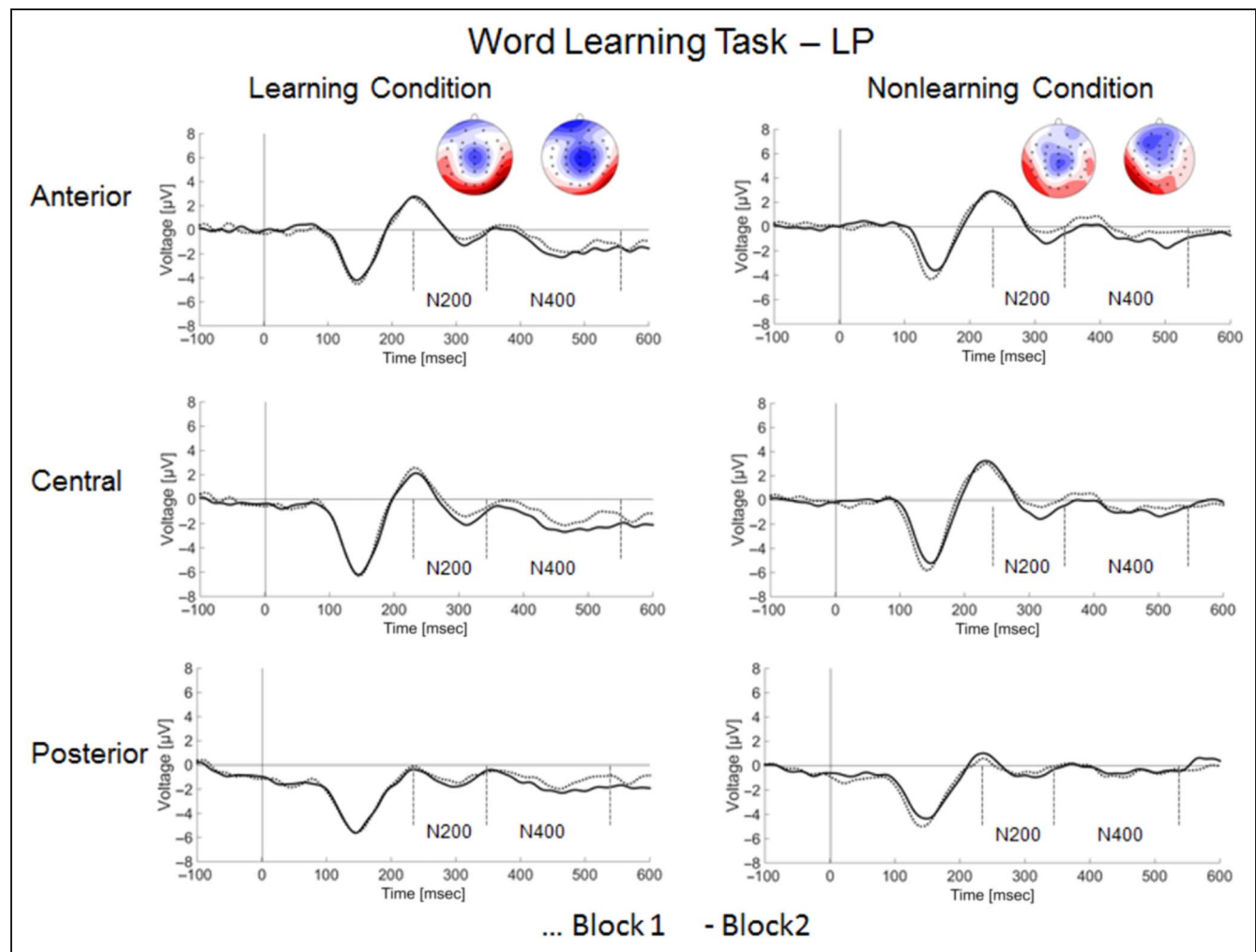


Figure 5. The EEG traces of the LP group are shown separately for the learning (left side) and nonlearning (right side) conditions at anterior (first row), central (second row), and posterior (third row) electrodes. Dotted line = first block, solid line = second block. The topographical maps depict voltage differences between the two blocks in the N200 and N400 time windows.

Table 1. Summary of the Significant Results of the Phonetic-based Word Learning Task

| Component | ANOVA | Results | F Value | p Value | Origin of the Effects |
|-----------|--------------------------------|---|---------|---------|---|
| N200 | $2 \times 2 \times 2 \times 3$ | Block | 9.874 | .003 | Block 2 > Block 1 |
| | | ROI | 4.194 | .038 | Central > Anterior |
| | | Condition \times ROI | 4.134 | .037 | Posterior: Learning > Nonlearning |
| | | Block \times ROI | 3.784 | .05 | Anterior/Central: Block 2 > Block 1 |
| N400 | $2 \times 2 \times 2 \times 3$ | Condition | 9.454 | .004 | Learning > Nonlearning |
| | | Block | 5.165 | .029 | Block 2 > Block 1 |
| | | Block \times ROI | 4.051 | .044 | Anterior/Central: Block 2 > Block 1 |
| | | Condition \times Block \times ROI | 3.746 | .017 | Block 2, Central/Posterior: Learning > Nonlearning |
| | | Group \times Condition \times Block \times ROI | 5.098 | .03 | HP, Nonlearning, Anterior/Central: Block 2 > Block 1 |

origin of the Block \times ROI interaction showed that, in Block 2, N400 responses at anterior and central electrodes were increased compared to Block 1 (anterior: $t(38) = 2.762, p = .009$; central: $t(38) = 2.523, p = .016$; posterior: $t(38) = 0.972, p = .337$; Figure 3). Otherwise, the significant Condition \times Block \times ROI interaction was further inspected by separate ANOVAs for the two blocks. These analyses yielded a Condition \times ROI interaction in Block 2 only, $F(1, 37) = 5.852, p = .012$, and further t tests indicated that N400 amplitudes were larger in the learning compared to the nonlearning condition at central and posterior electrodes (anterior: $t(38) = -0.922, p = .363$; central: $t(38) = -2.52, p = .016$; posterior: $t(38) = -3.507,$

$p = .001$; Figure 3). Finally, the Group \times Condition \times Block \times ROI interaction was further inspected by separate ANOVAs for the two groups. Because the Condition \times Block \times ROI interaction reached significance in the HP group only (HP: $F(1, 18) = 6.047, p = .024$; LP: $F(1, 19) = 0.796, p = .383$), we performed additional Block \times ROI ANOVAs, separately for the two conditions. This statistical procedure reached a significant Block \times ROI interaction in the nonlearning condition only (learning: $F(1, 18) = .073, p = .791$; nonlearning: $F(1, 18) = 10.019, p = .005$), and t tests within the HP group showed that, in Block 2 of the nonlearning condition, the N400 component was more negative than in Block 1 at anterior and central electrodes

Figure 6. Difference waves of the phonetic-based word learning task are plotted separately for the two groups (HPs = red line and LPs = blue line), the two blocks (left and right columns), and the anterior (first row), central (second row), and posterior (third row) electrodes.

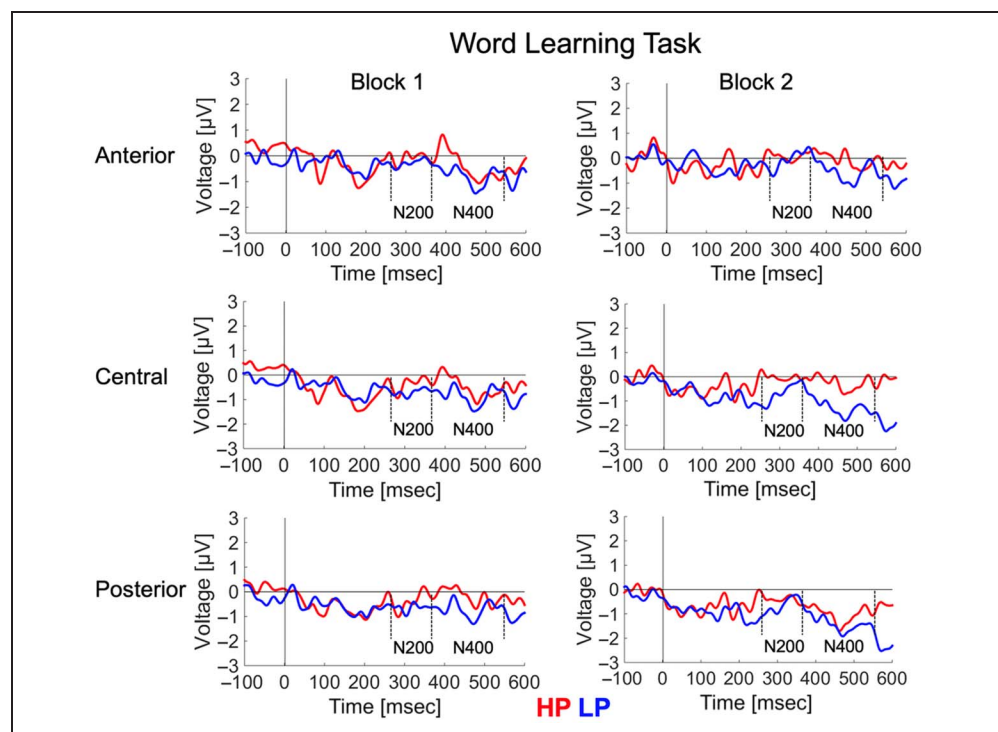


Table 2. Summary of the Significant Results of the Phonetic-based Word Learning Task; Evaluation of Difference Values Obtained by Subtracting the Nonlearning from the Learning Condition

| <i>Component</i> | <i>ANOVA</i> | <i>Results</i> | <i>F Value</i> | <i>p Value</i> | <i>Origin of the Effects</i> |
|------------------|--------------|---------------------|----------------|----------------|---|
| N200 | 2 × 2 × 3 | ROI | 4.134 | .037 | Central/Posterior > Anterior |
| N400 | 2 × 2 × 3 | Block × ROI | 5.395 | .017 | Block 2: Central/Posterior > Anterior; Posterior > Central |
| | | Group × Block × ROI | 5.098 | .03 | HP, Block 2: Posterior > Central LP, Block 2: Central > Anterior |

(anterior: $t(18) = 3.109, p = .006$; central: $t(18) = 3.137, p = .006$; posterior: $t(18) = -0.193, p = .849$; Figures 4 and 5 and Table 1).

Finally, in additional statistical analyses, we computed $2 \times 2 \times 3$ ANOVAs (two groups, two blocks, and three ROIs) with the difference values obtained by subtracting the nonlearning from the learning condition in the time windows corresponding to the N200 and N400 components. This more conservative approach was used to correct for learning-unspecific task-related EEG manifestations. The evaluation of the N200 component yielded a main effect of ROI, $F(2, 74) = 4.134, p = .037$; partial $\eta^2 = .101$, that was related to increased N200 negativity at central and posterior electrodes compared to anterior ones (anterior-central: $t(38) = 2.662, p = .011$; anterior-posterior: $t(38) = 2.130, p = .04$; central-posterior: $t(38) = 0.79, p = .434$; Figure 6). Otherwise, the N400 analysis revealed significant Block × ROI, $F(2, 74) = 5.395, p = .017$; partial $\eta^2 = .127$, and Group × Block × ROI (quadratic, $F(2, 74) = 5.098, p = .03$; partial $\eta^2 = .121$) interactions. The Block × ROI interaction was further inspected by separate t tests for the two blocks. These post hoc comparisons reached significance in Block 2 (anterior-central: $t(38) = 2.149, p = .038$; anterior-posterior: $t(38) = 2.619, p = .013$; central-posterior: $t(38) = 2.103, p = .042$) but not in Block 1 (anterior-central: $t(38) = 0.42, p = .677$; anterior-posterior: $t(38) = -0.378, p = .708$; central-posterior: $t(38) = -1.319, p = .195$), with increased N400 negativity at central and posterior electrodes compared to anterior ones and increased negativity at posterior compared to central electrodes (Figure 6). Finally, separate ANOVAs for the two blocks targeting at elucidating the Group × Block × ROI interaction showed a Group × ROI interaction only in Block 2 (Block 1: $F(2, 74) = 0.188, p = .667$; Block 2: $F(2, 74) = 6.094, p = .018$). Additional post hoc t tests for the two groups revealed that, in HPs, the N400 was more negative at the posterior compared to the central ROI (anterior-central: $t(18) = 0.691, p = .498$; anterior-posterior: $t(18) = 1.814, p = .086$; central-posterior: $t(18) = 2.6, p = .018$; Figure 6), whereas LPs demonstrated increased negativity at central compared to anterior electrodes (anterior-central: $t(19) = 2.271, p = .035$; anterior-posterior: $t(19) = 1.848, p = .08$; central-posterior: $t(19) = 0.444, p = .662$; Figure 6 and Table 2).

DISCUSSION

General Discussion

In this EEG study, we recruited a sample of 40 participants varying in their degree of musicality and divided them into two groups, HPs and LPs, based on the median split of the word learning performance. During the EEG experiment, the participants performed a phonetic-based word learning task that also included a nonlearning control condition. Furthermore, we collected behavioral data (d' values) while the participants completed an attentive listening and a phonetic discrimination task. The purpose of the study was to examine the specific and unspecific electrophysiological correlates of phonetic-based word learning and to test relationships between word learning performance and auditory attention functions, verbal memory capacity, and phonetic discrimination skills. Furthermore, we correlated musicality with phonetic-based word learning performance, auditory attention, and phonetic discrimination skills.

The evaluation of the psychometric data showed that HPs demonstrated better verbal memory functions than LPs. Importantly, during the phonetic-based word-meaning acquisition process, we revealed a topographical dissociation of N200/N400 manifestations between the learning and the nonlearning conditions. In particular, learning-specific N200/N400 responses emerged at central-posterior electrodes, whereas unspecific ones, possibly reflecting task-related cognitive engagement, were restricted to anterior scalp sites. Furthermore, in HPs, the N400 component was more negative in Block 2 than in Block 1 of the nonlearning condition at anterior and central electrodes. However, because it was unclear whether this effect was associated with the recruitment of additional cognitive resources or simply reflected the awareness of not having learned (mismatch effect), we conducted more conservative analyses with difference values obtained by subtracting the nonlearning from the learning condition. Notably, after having corrected the EEG traces for learning-unspecific effects, we noticed a performance-dependent N400 gradient along the anterior-posterior topographical axis, with maximal amplitudes at posterior electrodes in HPs, and more anterior maxima at central scalp sites in LPs. Finally, behavior-

behavior correlations provided additional insights into the multifaceted dimensions of phonetic-based word learning. In particular, word learning performance was positively related to phonetic discrimination skills and verbal memory capacity, whereas musicality correlated with phonetic discrimination abilities.

Attentive Listening and Phonetic Discrimination Tasks

Tonic alertness is commonly assessed using tasks requiring the detection of rare occurrences of target stimuli (Hamm et al., 2015; Tan, Gross, & Uhlhaas, 2015). Accordingly, in the attentive listening task, the participants had to detect infrequent repetitions of the same Thai words that were also used in the phonetic discrimination and phonetic-based word learning tasks. A close correspondence between the stimulus materials in all three tasks bears the advantage of facilitating contingencies between auditory attention functions, phonetic discrimination skills, and phonetic-based word learning performance. Nevertheless, it is important to note that the attentive listening task also strongly relied on phonetic discrimination skills because the participants had to distinguish between same (repeated) and different stimuli. This perspective is further corroborated by the positive correlation between the behavioral data (d' values) of the attentive listening and phonetic discrimination tasks (two-tailed, $r = .331, p = .040$).

The attentive listening and phonetic discrimination tasks served primarily to test relationships between auditory attention and phonetic discrimination functions, word learning performance, and musicality. The correlation analyses yielded a positive association between musicality and phonetic discrimination skills. Furthermore, phonetic discrimination skills were predictive of word learning performance. The relationship between phonetic discrimination skills and musicality is compatible with previous studies showing relationships between musical practice and segmental speech processing (Nan et al., 2018; Elmer, Hanggi, Meyer, & Jancke, 2013; Kuhnis, Elmer, Meyer, & Jancke, 2013), and might indicate that participants with a more sensitive “musical ear” are better able to discriminate Thai words than participants with a lower degree of musicality. Such a perspective would also be in agreement with previous studies that used the same Thai words as in the present work (Ditinger et al., 2016, 2017, 2019) and demonstrated that French musicians detected the difference between native and nonnative phonemes better than French nonmusicians (Ditinger, D’Imperio, & Besson, 2018). However, it is noteworthy to mention that, in a recent EEG study, Ditinger et al. (Ditinger, Korke, & Besson, 2021) used different sets of words and picture–word associations to neutralize the perceptual and associative learning advantages previously observed in musicians during word learning tasks (Ditinger et al., 2016, 2017, 2019). In this context, the authors could show that, in the absence of

between-group differences in auditory perception and discrimination, musicians are no longer advantaged in word learning. Hence, the positive correlation we revealed between musicality and phonetic discrimination skills might indicate that musical skills only facilitate word discrimination and word learning when the stimuli are difficult to differentiate. Finally, because the stimuli used in the phonetic-based word learning experiment were not familiar to the participants and were difficult to discriminate, we also expected a positive relationship between phonetic discrimination skills and word learning performance. This relationship testifies that the discrimination of the auditory stimuli constituted a “bottleneck” for word learning and that the distinction of the unfamiliar phonetic contrasts used in this study was a fundamental prerequisite for associating the new words with a meaning.

Phonetic-based Word Learning: Learning-specific and Unspecific EEG Effects

Based on the median split procedure applied, HPs not only outperformed LPs in the phonetic-based word learning task but were also more aware of the words they did not learn in the nonlearning condition. In this context, it is important to note that the participants were not informed about the inclusion of a nonlearning condition. The N200 component showed a maximum expression at central electrodes (main effect of ROI) and generally increased across the two blocks (main effect of Block), especially at anterior and central electrodes (Block \times ROI interaction). Furthermore, the examination of canonical N400 waveforms showed that this component was overall more pronounced in the learning compared to the nonlearning condition (main effect of Condition), however, without exhibiting a topographical specificity along the anterior–posterior axis. Such a topographically diffuse main effect of Condition confirmed the established sensitivity of the N400 component to lexical–semantic information (Kutas & Federmeier, 2011; Patel & Azzam, 2005; Brown & Hagoort, 1993). Notably, similar to the N200 ERP component, the N400 also increased in Block 2 compared to Block 1 (main effect of Block), irrespective of condition, and this effect was particularly pronounced at anterior and central electrodes (Block \times ROI interaction). The increased N200 and N400 amplitudes we noticed in the second compared to the first block at anterior and central electrodes are fully in line with the previous results of Ditinger et al. who used the same Thai words as in this study (Ditinger et al., 2016, 2017, 2019). However, because these anterior-central N200 and N400 distribution patterns were not differentially modulated by the two conditions, they are thought to reflect word learning-unspecific effects. This result is particularly interesting in that it enables a better understanding of the functional meaning of anteriorly distributed negativities in the context of word learning tasks. In particular, based on the assumption that in the nonlearning condition participants

tried to associate pictures with words in a similar way as they did in the learning condition, anterior N200 and N400 effects were interpreted as reflecting increased allocation of processing resources to maintain novel information in working memory (Hagoort, 2014; Patel & Azzam, 2005) and to form episodic memory traces of picture–word associations (Ditinger et al., 2016). Furthermore, anterior N200 and N400 manifestations could also reflect domain-general cognitive control mechanisms involved in the early stages of learning even when learning is not possible (Rodriguez-Fornells et al., 2009; Mestres-Misse et al., 2007).

In this study, we consistently revealed larger N200 and N400 amplitudes at posterior scalp sites in the learning compared to the nonlearning condition. These results suggest a sensitivity of posterior electrodes to the word–meaning acquisition process and are thought to reflect a learning-specific engagement of neural resources. Furthermore, such a sensitivity of posterior electrodes to learning mechanisms was also confirmed by the evaluation of the difference values obtained by subtracting the nonlearning from the learning condition. In particular, the additional analyses that were corrected for learning-unspecific effects (Figure 6) clearly demonstrated increased N200-related amplitudes at central-posterior electrodes compared to anterior ones. In addition, in the second block of the word learning task, we noticed an increased N400-related negativity at central and posterior compared to anterior electrodes, with a maximum expression at posterior scalp sites. Because the central-posterior N200-related negativity did not specifically increase across the two blocks of the phonetic-based word learning condition, we conclude that this electrophysiological pattern did not mirror incremental learning but rather the demands placed on early lexical selection and initial word–form recognition (van den Brink et al., 2001). Otherwise, the N400-related results that were corrected for learning-unspecific effects corroborate previous findings showing that especially a posterior distribution of this ERP component can be considered as a sensitive marker of information integration into lexical–semantic memory (Kutas & Federmeier, 2011; Patel & Azzam, 2005).

In conclusion, anterior N200/N400 manifestations indicate learning-unspecific effects that possibly mirror an increased allocation of processing resources or cognitive control mechanisms, which are needed to maintain novel information in working memory (Hagoort, 2014; Patel & Azzam, 2005) and to form episodic memory traces (Ditinger et al., 2016). In contrast, posterior N200/N400 distributions can be used as reliable markers of word learning and reflect the integration of novel words into lexical–semantic memory.

Phonetic-based Word Learning: Performance-dependent EEG Effects

Canonical N400 analyses revealed a Group \times Condition \times Block \times ROI interaction effect that was related to

increased N400 amplitudes in HPs at anterior and central electrodes in the second compared to the first block of the nonlearning condition. Such a condition-specific effect could lead to the interpretations, among others, that HPs mobilized additional cognitive resources to try to acquire the meaning of the words of the nonlearning condition, or that they were aware that they did not learn them. However, the evaluation of N400-related amplitudes that were corrected for learning-unspecific effects yielded a Group \times Block \times ROI interaction that was driven by increased N400 amplitudes at posterior compared to central electrodes in HPs, whereas LPs were characterized by increased amplitudes at central compared to anterior electrodes. These results are in line with the existing N400 literature on word learning (Borovsky, Elman, & Kutas, 2012; Borovsky et al., 2010; Mestres-Misse et al., 2007), and support the view that the more posterior N400-related manifestations in HPs can be ascribed to a facilitated access to lexical–semantic memory (Kutas & Federmeier, 2011; Patel & Azzam, 2005). Furthermore, it is noteworthy to mention that phonetic-based word learning performance was positively related to verbal memory and that HPs had a better verbal memory capacity than LPs. These results are interesting in that they further contribute to a better understanding of the multifaceted dimensions of phonetic-based word learning. In particular, our results suggest that learning the meaning of new words based on picture–word associations is dependent upon individual verbal memory span. Because both the VLMT test and the phonetic-based word learning task used in this study rely on verbal working memory functions, the relationship we found between the two variables could be expected. In fact, verbal working memory has traditionally been assigned to frontoparietal brain regions (Emch, von Bastian, & Koch, 2019; Koelsch et al., 2009; Veltman, Rombouts, & Dolan, 2003) and is known to be related to several other cognitive functions and language abilities (Linck, Osthus, Koeth, & Bunting, 2014; Montgomery, 2003; Daneman & Merikle, 1996). Hence, it is not surprising that the ability to hold verbal information in a short-term buffer is helpful for learning unfamiliar phonemes as well as their associations with conceptual objects (Hartley & Houghton, 1996).

Conclusions

In the present work, we proposed a methodological framework of phonetic-based word learning that is helpful to disentangle specific from unspecific word learning effects and to reduce the risk of misleading ERP interpretations. Furthermore, we provided additional insights into the performance-dependent electrophysiological correlates of phonetic-based word learning and scrutinized the possible role of verbal memory capacity, auditory attention functions, phonetic discrimination skills, and musicality in phonetic-based word learning.

Acknowledgments

This research was supported by the Swiss National Science Foundation (grant no. 320030_163149 to L. J.). We would like to thank Dorothea Weniger for her help in the selection of the stimulus material.

Reprint requests should be sent to Stefan Elmer, Institute of Psychology, Division Neuropsychology, University of Zurich, Binzmühlestrasse 14/25, 8050 Zurich, Switzerland, or via e-mail: s.elmer@psychologie.uzh.ch.

Author Contributions

Stefan Elmer: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Writing—Original draft. Eva Dittinger: Conceptualization. Julia Brocchetto: Data curation; Investigation. Clément François: Conceptualization; Writing—Review & editing. Mireille Besson: Conceptualization; Writing—Review & editing. Lutz Jäncke: Funding acquisition; Writing—Review & editing. Antoni Rodríguez-Fornells: Conceptualization; Writing—Review & editing.

Code and Data Availability

The EEG data of this study are available from the first author upon reasonable request.

Funding Information

Lutz Jäncke, Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung (<http://dx.doi.org/10.13039/501100001711>), grant number: 320030_163149.

Diversity in Citation Practices

A retrospective analysis of the citations in every article published in this journal from 2010 to 2020 has revealed a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience (JoCN)* during this period were $M(\text{an})/M = .408$, $W(\text{oman})/M = .335$, $M/W = .108$, and $W/W = .149$, the comparable proportions for the articles that these authorship teams cited were $M/M = .579$, $W/M = .243$, $M/W = .102$, and $W/W = .076$ (Fulvio et al., *JoCN*, 33:1, pp. 3–7). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance.

REFERENCES

Annett, M. (1970). A classification of hand preference by association analysis. *British Journal of Psychology*, 61, 303–321. <https://doi.org/10.1111/j.2044-8295.1970.tb01248.x>
 Bahramlou, K., & Esmaeili, A. (2019). The effects of vocabulary enhancement exercises and group dynamic assessment on word learning through lexical inferencing. *Journal of*

Psycholinguistic Research, 48, 889–901. <https://doi.org/10.1007/s10936-019-09638-x>, PubMed: 30903423
 Bakker, I., Takashima, A., van Hell, J. G., Janzen, G., & McQueen, J. M. (2015). Tracking lexical consolidation with ERPs: Lexical and semantic-priming effects on N400 and LPC responses to newly-learned words. *Neuropsychologia*, 79, 33–41. <https://doi.org/10.1016/j.neuropsychologia.2015.10.020>, PubMed: 26476370
 Bakker-Marshall, I., Takashima, A., Schoffelen, J. M., van Hell, J. G., Janzen, G., & McQueen, J. M. (2018). Theta-band oscillations in the middle temporal gyrus reflect novel word consolidation. *Journal of Cognitive Neuroscience*, 30, 621–633. https://doi.org/10.1162/jocn_a_01240, PubMed: 29393716
 Barbarotto, R., Laiacona, M., Macchi, V., & Capitani, E. (2002). Picture reality decision, semantic categories and gender. A new set of pictures, with norms and an experimental study. *Neuropsychologia*, 40, 1637–1653. [https://doi.org/10.1016/S0028-3932\(02\)00029-5](https://doi.org/10.1016/S0028-3932(02)00029-5), PubMed: 11992652
 Batterink, L., & Neville, H. (2011). Implicit and explicit mechanisms of word learning in a narrative context: An event-related potential study. *Journal of Cognitive Neuroscience*, 23, 3181–3196. https://doi.org/10.1162/jocn_a_00013, PubMed: 21452941
 Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: common processing, attention, and memory. *Frontiers in Psychology*, 2, 94. <https://doi.org/10.3389/fpsyg.2011.00094>, PubMed: 21738519
 Borovsky, A., Elman, J. L., & Kutas, M. (2012). Once is enough: N400 indexes semantic integration of novel word meanings from a single exposure in context. *Language Learning and Development*, 8, 278–302. <https://doi.org/10.1080/15475441.2011.614893>, PubMed: 23125559
 Borovsky, A., Kutas, M., & Elman, J. (2010). Learning to use words: Event-related potentials index single-shot contextual word learning. *Cognition*, 116, 289–296. <https://doi.org/10.1016/j.cognition.2010.05.004>, PubMed: 20621846
 Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, 5, 34–44. <https://doi.org/10.1162/jocn.1993.5.1.34>, PubMed: 23972118
 Connolly, J. F., & Phillips, N. A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience*, 6, 256–266. <https://doi.org/10.1162/jocn.1994.6.3.256>, PubMed: 23964975
 Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, 3, 422–433. <https://doi.org/10.3758/BF03214546>, PubMed: 24213976
 de Diego-Balaguer, R., Martinez-Alvarez, A., & Pons, F. (2016). Temporal attention as a scaffold for language development. *Frontiers in Psychology*, 7, 44. <https://doi.org/10.3389/fpsyg.2016.00044>, PubMed: 26869953
 Dimoska, A., Johnstone, S. J., & Barry, R. J. (2006). The auditory-evoked N2 and P3 components in the stop-signal task: Indices of inhibition, response-conflict or error-detection? *Brain and Cognition*, 62, 98–112. <https://doi.org/10.1016/j.bandc.2006.03.011>, PubMed: 16814442
 Dittinger, E., Barbaroux, M., D'Imperio, M., Jancke, L., Elmer, S., & Besson, M. (2016). Professional music training and novel word learning: From faster semantic encoding to longer-lasting word representations. *Journal of Cognitive Neuroscience*, 28, 1584–1602. https://doi.org/10.1162/jocn_a_00997, PubMed: 27315272
 Dittinger, E., Chobert, J., Ziegler, J. C., & Besson, M. (2017). Fast brain plasticity during word learning in musically-trained children. *Frontiers in Human Neuroscience*, 11, 233. <https://doi.org/10.3389/fnhum.2017.00233>, PubMed: 28553213

- Dittinger, E., D'Imperio, M., & Besson, M. (2018). Enhanced neural and behavioural processing of a nonnative phonemic contrast in professional musicians. *European Journal of Neuroscience*, *47*, 1504–1516. <https://doi.org/10.1111/ejn.13939>, PubMed: 29753304
- Dittinger, E., Korke, B., & Besson, M. (2021). Evidence for enhanced long-term memory in professional musicians and its contribution to novel word learning. *Journal of Cognitive Neuroscience*, *33*, 662–682. https://doi.org/10.1162/jocn_a_01670, PubMed: 33378241
- Dittinger, E., Scherer, J., Jancke, L., Besson, M., & Elmer, S. (2019). Testing the influence of musical expertise on novel word learning across the lifespan using a cross-sectional approach in children, young adults and older adults. *Brain and Language*, *198*, 104678. <https://doi.org/10.1016/j.bandl.2019.104678>, PubMed: 31450024
- Dobel, C., Lagemann, L., & Zwitserlood, P. (2009). Non-native phonemes in adult word learning: evidence from the N400m. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *364*, 3697–3709. <https://doi.org/10.1098/rstb.2009.0158>, PubMed: 19933141
- Eiteljoerge, S. F. V., Adam, M., Elsner, B., & Mani, N. (2019). Consistency of co-occurring actions influences young children's word learning. *Royal Society Open Science*, *6*, 190097. <https://doi.org/10.1098/rsos.190097>, PubMed: 31598229
- Elmer, S., Greber, M., Pushparaj, A., Kuhnis, J., & Jancke, L. (2017). Faster native vowel discrimination learning in musicians is mediated by an optimization of mnemonic functions. *Neuropsychologia*, *104*, 64–75. <https://doi.org/10.1016/j.neuropsychologia.2017.08.001>, PubMed: 28780308
- Elmer, S., Hanggi, J., Meyer, M., & Jancke, L. (2013). Increased cortical surface area of the left planum temporale in musicians facilitates the categorization of phonetic and temporal speech sounds. *Cortex*, *49*, 2812–2821. <https://doi.org/10.1016/j.cortex.2013.03.007>, PubMed: 23628644
- Elmer, S., Meyer, M., & Jancke, L. (2012). Neurofunctional and behavioral correlates of phonetic and temporal categorization in musically trained and untrained subjects. *Cerebral Cortex*, *22*, 650–658. <https://doi.org/10.1093/cercor/bhr142>, PubMed: 21680844
- Emch, M., von Bastian, C. C., & Koch, K. (2019). Neural correlates of verbal working memory: An fMRI meta-analysis. *Frontiers in Human Neuroscience*, *13*, 180. <https://doi.org/10.3389/fnhum.2019.00180>
- Enriquez-Geppert, S., Konrad, C., Pantev, C., & Huster, R. J. (2010). Conflict and inhibition differentially affect the N200/P300 complex in a combined go/nogo and stop-signal task. *Neuroimage*, *51*, 877–887. <https://doi.org/10.1016/j.neuroimage.2010.02.043>, PubMed: 20188191
- Francois, C., Cunillera, T., Garcia, E., Laine, M., & Rodriguez-Fomells, A. (2017). Neurophysiological evidence for the interplay of speech segmentation and word-referent mapping during novel word learning. *Neuropsychologia*, *98*, 56–67. <https://doi.org/10.1016/j.neuropsychologia.2016.10.006>, PubMed: 27732869
- Friedrich, M., & Friederici, A. D. (2008). Neurophysiological correlates of online word learning in 14-month-old infants. *NeuroReport*, *19*, 1757–1761. <https://doi.org/10.1097/WNR.0b013e328318f014>, PubMed: 18955904
- Frishkoff, G. A., Perfetti, C. A., & Collins-Thompson, K. (2010). Lexical quality in the brain: ERP evidence for robust word learning from context. *Developmental Neuropsychology*, *35*, 376–403. <https://doi.org/10.1080/87565641.2010.480915>, PubMed: 20614356
- Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Auditory attention—Focusing the searchlight on sound. *Current Opinion in Neurobiology*, *17*, 437–455. <https://doi.org/10.1016/j.conb.2007.07.011>, PubMed: 17714933
- Garagnani, M., Wennekers, T., & Pulvermuller, F. (2008). A neuroanatomically grounded Hebbian-learning model of attention-language interactions in the human brain. *European Journal of Neuroscience*, *27*, 492–513. <https://doi.org/10.1111/j.1460-9568.2008.06015.x>, PubMed: 18215243
- Getzmann, S., Jasny, J., & Falkenstein, M. (2017). Switching of auditory attention in “cocktail-party” listening: ERP evidence of cueing effects in younger and older adults. *Brain and Cognition*, *111*, 1–12. <https://doi.org/10.1016/j.bandc.2016.09.006>, PubMed: 27814564
- Gordon, E. E. (1989). *Manual for the advanced measures of music education*. Chicago: G. I. A. Publications.
- Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca's region and beyond. *Current Opinion in Neurobiology*, *28*, 136–141. <https://doi.org/10.1016/j.conb.2014.07.013>, PubMed: 25062474
- Hamm, J. P., Bobilev, A. M., Hayrynen, L. K., Hudgens-Haney, M. E., Oliver, W. T., Parker, D. A., et al. (2015). Stimulus train duration but not attention moderates gamma-band entrainment abnormalities in schizophrenia. *Schizophrenia Research*, *165*, 97–102. <https://doi.org/10.1016/j.schres.2015.02.016>, PubMed: 25868936
- Hartley, T., & Houghton, G. (1996). A linguistically constrained model of short-term memory for nonwords. *Journal of Memory and Language*, *35*, 1–31. <https://doi.org/10.1006/jmla.1996.0001>
- Helmstaedter, C. L. M., & Lux, S. (2001). Verbaler Lern- und Merkfähigkeitstest (VLMT). *Beltz, Göttingen (2001)*.
- Junge, C., Cutler, A., & Hagoort, P. (2012). Electrophysiological evidence of early word learning. *Neuropsychologia*, *50*, 3702–3712. <https://doi.org/10.1016/j.neuropsychologia.2012.10.012>, PubMed: 23108241
- Kim, A. E., Oines, L., & Miyake, A. (2018). Individual differences in verbal working memory underlie a tradeoff between semantic and structural processing difficulty during language comprehension: An ERP Investigation. *Journal of Experimental Psychology: Learning Memory and Cognition*, *44*, 406–420. <https://doi.org/10.1037/xlm0000457>, PubMed: 28933902
- Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Muller, K., & Gruber, O. (2009). Functional architecture of verbal and tonal working memory: An fMRI study. *Human Brain Mapping*, *30*, 859–873. <https://doi.org/10.1002/hbm.20550>, PubMed: 18330870
- Kuhnis, J., Elmer, S., Meyer, M., & Jancke, L. (2013). The encoding of vowels and temporal speech cues in the auditory cortex of professional musicians: An EEG study. *Neuropsychologia*, *51*, 1608–1618. <https://doi.org/10.1016/j.neuropsychologia.2013.04.007>, PubMed: 23664833
- 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>, PubMed: 20809790
- Langer, N., von Bastian, C. C., Wirz, H., Oberauer, K., & Jancke, L. (2013). The effects of working memory training on functional brain network efficiency. *Cortex*, *49*, 2424–2438. <https://doi.org/10.1016/j.cortex.2013.01.008>, PubMed: 23489778
- Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience*, *9*, 920–933. <https://doi.org/10.1038/nrn2532>, PubMed: 19020511
- Lehrl, S. G. A., Blaha, L., & Fischer, B. (1992). *Kurztest fu r allgemeine intelligenz (KAI)*. Ebersberg: Vless.
- Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin &*

- Review*, 21, 861–883. <https://doi.org/10.3758/s13423-013-0565-2>, PubMed: 24366687
- Maess, B., Friederici, A. D., Damian, M., Meyer, A. S., & Levelt, W. J. (2002). Semantic category interference in overt picture naming: Sharpening current density localization by PCA. *Journal of Cognitive Neuroscience*, 14, 455–462. <https://doi.org/10.1162/089892902317361967>, PubMed: 11970804
- Mestres-Misse, A., Rodriguez-Fornells, A., & Munte, T. F. (2007). Watching the brain during meaning acquisition. *Cerebral Cortex*, 17, 1858–1866. <https://doi.org/10.1093/cercor/bhl094>, PubMed: 17056648
- Michie, P. T., Bearpark, H. M., Crawford, J. M., & Glue, L. C. (1990). The nature of selective attention effects on auditory event-related potentials. *Biological Psychology*, 30, 219–250. [https://doi.org/10.1016/0301-0511\(90\)90141-I](https://doi.org/10.1016/0301-0511(90)90141-I), PubMed: 2282370
- Montgomery, J. W. (2003). Working memory and comprehension in children with specific language impairment: What we know so far. *Journal of Communication Disorders*, 36, 221–231. [https://doi.org/10.1016/S0021-9924\(03\)00021-2](https://doi.org/10.1016/S0021-9924(03)00021-2), PubMed: 12742669
- Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological Science*, 22, 1425–1433. <https://doi.org/10.1177/0956797611416999>, PubMed: 21969312
- Nan, Y., Liu, L., Geiser, E., Shu, H., Gong, C. C., Dong, Q., et al. (2018). Piano training enhances the neural processing of pitch and improves speech perception in Mandarin-speaking children. *Proceedings of the National Academy of Sciences, U.S.A.*, 115, E6630–E6639. <https://doi.org/10.1073/pnas.1808412115>, PubMed: 29941577
- Patel, S. H., & Azzam, P. N. (2005). Characterization of N200 and P300: Selected studies of the event-related potential. *International Journal of Medical Sciences*, 2, 147–154. <https://doi.org/10.7150/ijms.2.147>, PubMed: 16239953
- Rasamimanana, M., Barbaroux, M., Cole, P., & Besson, M. (2020). Semantic compensation and novel word learning in university students with dyslexia. *Neuropsychologia*, 139, 107358. <https://doi.org/10.1016/j.neuropsychologia.2020.107358>, PubMed: 31978401
- Rodriguez-Fornells, A., Cunillera, T., Mestres-Misse, A., & de Diego-Balaguer, R. (2009). Neurophysiological mechanisms involved in language learning in adults. *Philosophical Transactions of the Royal Society of London: Series B, Biological Sciences*, 364, 3711–3735. <https://doi.org/10.1098/rstb.2009.0130>, PubMed: 19933142
- Rueda-Delgado, L. M., Heise, K. F., Daffertshofer, A., Mantini, D., & Swinnen, S. P. (2019). Age-related differences in neural spectral power during motor learning. *Neurobiology of Aging*, 77, 44–57. <https://doi.org/10.1016/j.neurobiolaging.2018.12.013>, PubMed: 30776651
- Schroger, E. (1993). Event-related potentials to auditory-stimuli following transient shifts of spatial attention in a go nogo task. *Biological Psychology*, 36, 183–207. [https://doi.org/10.1016/0301-0511\(93\)90017-3](https://doi.org/10.1016/0301-0511(93)90017-3), PubMed: 8260565
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, 31, 137–149. <https://doi.org/10.3758/BF03207704>
- Strait, D. L., Slater, J., O'Connell, S., & Kraus, N. (2015). Music training relates to the development of neural mechanisms of selective auditory attention. *Developmental Cognitive Neuroscience*, 12, 94–104. <https://doi.org/10.1016/j.dcn.2015.01.001>, PubMed: 25660985
- Takashima, A., Bakker, L., van Hell, J. G., Janzen, G., & McQueen, J. M. (2017). Interaction between episodic and semantic memory networks in the acquisition and consolidation of novel spoken words. *Brain and Language*, 167, 44–60. <https://doi.org/10.1016/j.bandl.2016.05.009>, PubMed: 27291335
- Tan, H. R. M., Gross, J., & Uhlhaas, P. J. (2015). MEG-measured auditory steady-state oscillations show high test–retest reliability: A sensor and source-space analysis. *Neuroimage*, 122, 417–426. <https://doi.org/10.1016/j.neuroimage.2015.07.055>, PubMed: 26216274
- van den Brink, D., Brown, C. M., & Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *Journal of Cognitive Neuroscience*, 13, 967–985. <https://doi.org/10.1162/089892901753165872>, PubMed: 11595099
- Veltman, D. J., Rombouts, S. A., & Dolan, R. J. (2003). Maintenance versus manipulation in verbal working memory revisited: An fMRI study. *Neuroimage*, 18, 247–256. [https://doi.org/10.1016/S1053-8119\(02\)00049-6](https://doi.org/10.1016/S1053-8119(02)00049-6), PubMed: 12595179
- von Koss Torkildsen, J., Svangstu, J. M., Hansen, H. F., Smith, L., Simonsen, H. G., Moen, I., et al. (2008). Productive vocabulary size predicts event-related potential correlates of fast mapping in 20-month-olds. *Journal of Cognitive Neuroscience*, 20, 1266–1282. <https://doi.org/10.1162/jocn.2008.20087>, PubMed: 18284350
- Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., et al. (1998). Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. *Science*, 281, 1188–1191. <https://doi.org/10.1126/science.281.5380.1188>, PubMed: 9712582
- Wang, H., Zhang, G., & Liu, B. (2017). Influence of auditory spatial attention on cross-modal semantic priming effect: Evidence from N400 effect. *Experimental Brain Research*, 235, 331–339. <https://doi.org/10.1007/s00221-016-4792-4>, PubMed: 27722788
- Zendel, B. R., West, G. L., Belleville, S., & Peretz, I. (2019). Musical training improves the ability to understand speech-in-noise in older adults. *Neurobiology of Aging*, 81, 102–115. <https://doi.org/10.1016/j.neurobiolaging.2019.05.015>, PubMed: 31280114
- Zuk, J., Benjamin, C., Kenyon, A., & Gaab, N. (2014). Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PLoS One*, 9, e99868. <https://doi.org/10.1371/journal.pone.0099868>, PubMed: 24937544