



HAL
open science

Evidence for Enhanced Long-term Memory in Professional Musicians and Its Contribution to Novel Word Learning

Eva Dittinger, Betina Korka, Mireille R Besson

► **To cite this version:**

Eva Dittinger, Betina Korka, Mireille R Besson. Evidence for Enhanced Long-term Memory in Professional Musicians and Its Contribution to Novel Word Learning. *Journal of Cognitive Neuroscience*, 2021, 33, pp.662 - 682. 10.1162/jocn_a_01670 . hal-03170730

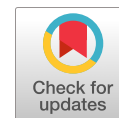
HAL Id: hal-03170730

<https://amu.hal.science/hal-03170730>

Submitted on 16 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Evidence for Enhanced Long-term Memory in Professional Musicians and Its Contribution to Novel Word Learning

Eva Dittinger^{1,2,3} , Betina Korka⁴, and Mireille Besson^{1,3}

Abstract

■ Previous studies evidenced transfer effects from professional music training to novel word learning. However, it is unclear whether such an advantage is driven by cascading, bottom-up effects from better auditory perception to semantic processing or by top-down influences from cognitive functions on perception. Moreover, the long-term effects of novel word learning remain an open issue. To address these questions, we used a word learning design, with four different sets of novel words, and we neutralized the potential perceptive and associative learning advantages in musicians. Under such conditions, we did not observe any advantage in musicians on the day of learning (Day 1 [D1]), at neither a behavioral nor an electrophysiological level; this suggests that the

previously reported advantages in musicians are likely to be related to bottom-up processes. Nevertheless, 1 month later (Day 30 [D30]) and for all types of novel words, the error increase from D1 to D30 was lower in musicians compared to nonmusicians. In addition, for the set of words that were perceptually difficult to discriminate, only musicians showed typical N400 effects over parietal sites on D30. These results demonstrate that music training improved long-term memory and that transfer effects from music training to word learning (i.e., semantic levels of speech processing) benefit from reinforced (long-term) memory functions. Finally, these findings highlight the positive impact of music training on the acquisition of foreign languages. ■

INTRODUCTION

Playing a musical instrument at a professional level is a multidimensional task that requires acute auditory perception, focused attention, the ability to maintain auditory information in short-term memory (STM) and long-term memory, and specific motor abilities. Typically, professional musicians start playing their instrument at a very young age and train for years, many hours per week. Accumulating evidence indicates that such an intensive training improves auditory perception and attention (Bidelman, Krishnan, & Gandour, 2011; Strait, Kraus, Parbery-Clark, & Ashley, 2010; Micheyl, Delhommeau, Perrot, & Oxenham, 2006; Tervaniemi, Castaneda, Knoll, & Uther, 2006; Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004; Kishon-Rabin, Amir, Vexler, & Zaltz, 2001; Koelsch, Schröger, & Tervaniemi, 1999; Trainor, Desjardins, & Rockel, 1999; Besson & Faïta, 1995; Spiegel & Watson, 1984) and strongly promotes brain plasticity (Münste, Altenmüller, & Jäncke, 2002). Most importantly, for this study, music-training-related advantages have been shown to transfer to the language domain as

well (for a review, see Asaridou & McQueen, 2013; Besson, Chobert, & Marie, 2011; Kraus & Chandrasekaran, 2010), having effects at the segmental (Linnavalli, Putkinen, Lipsanen, Huotilainen, & Tervaniemi, 2018; Intartaglia, White-Schwoch, Kraus, & Schön, 2017; Gordon, Fehd, & McCandliss, 2015; Bidelman, Weiss, Moreno, & Alain, 2014; Elmer, Meyer, & Jäncke, 2012; Chobert, Marie, François, Schön, & Besson, 2011), suprasegmental (Torppa, Faulkner, Laasonen, Lipsanen, & Sammler, 2020; Marie, Delogu, Lampis, Belardinelli, & Besson, 2011; Wong & Perrachione, 2007), and even syntactic (Fitzroy & Sanders, 2013) level of speech processing. The evidence also suggests that long-term music training improves second-language acquisition (children's listening comprehension, phonological and word knowledge of English: Yang, Ma, Gong, Hu, & Yao, 2014), auditory attention (Strait, Slater, O'Connell, & Kraus, 2015) and speech-in-noise perception (Slater et al., 2015), visual attention (Wang, Osher, & Reuter-Lorenz, 2015), working and verbal memory (Roden, Grube, Bongard, & Kreutz, 2014; George & Coch, 2011), executive functions (Habibi, Damasio, Ilari, Elliott Sachs, & Damasio, 2018; Jaschke, Honing, & Scherder, 2018; Zuk, Benjamin, Kenyon, & Gaab, 2014; for reviews, see Benz, Sellaro, Hommel, & Colzato, 2016; Moreno & Farzan, 2015; Moreno & Bidelman, 2014), and general intelligence (Schellenberg, 2004, 2011). However, it has also been argued that genetic factors are possibly mediating the relationships between music training and intelligence (Swaminathan, Schellenberg, & Khalil, 2017; Mosing, Madison, Pedersen, & Ullén, 2016) and that motivation may

¹Université Publique de France, CNRS & Aix-Marseille University, Laboratoire de Neurosciences Cognitives (LNC), ²Université Publique de France, CNRS & Aix-Marseille University, Laboratoire Parole et Langage (LPL), ³Institute for Language and Communication in the Brain, Aix-en-Provence, France, ⁴Cognitive and Biological Psychology, Institute of Psychology - Wilhelm Wundt, Leipzig University, Germany

be more important than music training itself so that children who find music rewarding are more likely to continue music training than children who do not (Schellenberg, Corrigall, Dys, & Malti, 2015).

Recently, by comparing musicians and nonmusicians in three independent samples (children, young adults, and older adults), Dittinger and colleagues (Dittinger, Scherer, Jäncke, Besson, & Elmer, 2019; Dittinger, Chobert, Ziegler, & Besson, 2017; Dittinger et al., 2016) showed that musical expertise influences the semantic level of speech processing. More specifically, the authors focused on novel word learning (a multidimensional task requiring both perceptive and cognitive functions) and conducted a series of experiments that comprised phonological categorization tasks, a picture-word learning phase, and a test phase. The test phase included a matching task (i.e., does the word match or not with the previously learned picture-word association?) and a semantic task (i.e., is the word semantically related or not to a novel picture?). First, both behavioral measures and ERPs indicated that the musicians outperformed the nonmusicians during the phonological categorization tasks based on aspiration and pitch, two nonnative phonetic contrasts for the French-speaking participants (Dittinger, D'Imperio, & Besson, 2018; Dittinger et al., 2016). Second, only the musicians showed a significant N400 increase in amplitude over fronto-central regions, from the first to second half of the picture-word learning phase (François, Cunillera, Garcia, Laine, & Rodriguez-Fornells, 2017; Batterink & Neville, 2011; McLaughlin, Osterhout, & Kim, 2004). Third, in both tasks of the test phase, musicians were characterized by significantly larger N400 amplitude to words that were unexpected compared to words that were expected, based on the previously learned picture-word associations. These N400 effects (i.e., the difference between unexpected minus expected words) were largest over centro-parietal sites, a scalp distribution that is comparable to the N400 effect observed for known words (Kutas, Van Petten, & Besson, 1988). Finally, these results at the electrophysiological level were accompanied by a higher level of performance in the semantic task for the musicians' group, compared to the nonmusicians' group. Importantly, the series of results described above for young adults (Dittinger et al., 2016) were similar in children (Dittinger et al., 2017) but less clear-cut in older adults (Dittinger et al., 2019).

Two main interpretations have been proposed to explain transfer effects from music training to novel word learning (Dittinger et al., 2016). First, according to the cascading interpretation, enhanced auditory perception drive transfer effects in musicians (Mankel & Bidelman, 2018) by facilitating the bottom-up processing stages involved in novel word learning (speech perception, building phonological, lexical, and semantic representations; Besson, Barbaroux, & Dittinger, 2017). Second, according to the multidimensional, top-down interpretation musicians do show not only enhanced speech perception but also improved attention (Strait et al., 2015), working memory and/or STM (Roden et al., 2014; Schulze, Dowling, & Tillmann, 2012; George

& Coch, 2011), and executive functions (Habibi et al., 2018; Moreno & Bidelman, 2014; Zuk et al., 2014), and these different functions may, in turn, all contribute to novel word learning. The aim of the present series of experiments was to disentangle these two interpretations and better understand why musicians have an advantage in novel word learning.

To this end, we performed a series of four experiments (see Figure 1). In Experiment 1 (E1), we used a similar design to Dittinger et al. (2016) that aimed at learning word-picture associations, using the same stimuli (monosyllabic Thai words) but with one important modification aimed at reducing the influence of auditory perception and associative learning. That is, rather than performing only one learning phase (to learn picture-word associations) followed by only one matching task (to test for associative learning), participants performed a variable number of learning-matching cycles, until each participant reached a level of 83% correct responses in the matching task. This ensured that all participants had learned the associations equally well before performing the semantic task. Similarly, in Experiment 2 (E2), the same learning-matching task cycles procedure was used, only with Finnish disyllabic words, which, by contrast to the monosyllabic Thai words, were easy to discriminate for all participants. This presumably further attenuated the potential influence of enhanced auditory speech perception in musicians (Strait et al., 2010, 2015; Bidelman et al., 2014). Consequently, if differences in novel word learning between musicians and nonmusicians are still to be found in E1 and E2, these would likely reflect the top-down influences of higher cognitive functions (i.e., attention; better integration of novel words into preexisting semantic networks and/or better memory).

Experiments 3 and 4 (E3 and E4) were designed as control experiments accounting for musicians' potential advantage in cross-modal integration of visual picture and auditory word associations (Bidelman, 2016; Paraskevopoulos, Kuchenbuch, Herholz, & Pantev, 2012). In E3, disyllabic Finnish words were associated with nonobjects (i.e., objects with no semantic content), so that participants had to learn an association between two unknown items (i.e., neither the pictures nor the words had an established semantic meaning). By contrast, in E4, participants had to learn an association between two known items. That is, known French words were associated with known but noncorresponding pictures (e.g., the word "fromage" [cheese] was presented together with the picture of a hat). If professional musicians perform better than nonmusicians in E3 and E4, this would suggest that facilitation in cross-modal integration contributes to explaining their advantage over the nonmusicians in novel word learning.

After having established whether (1) semantic integration is faster and/or more efficient in musicians even without the advantage of better auditory perception (E1 and E2) and (2) better cross-modal integration in musicians contributes to novel word learning (E3 and E4), the third aim was to investigate the long-term effects of novel word learning. Participants performed the semantic tasks from

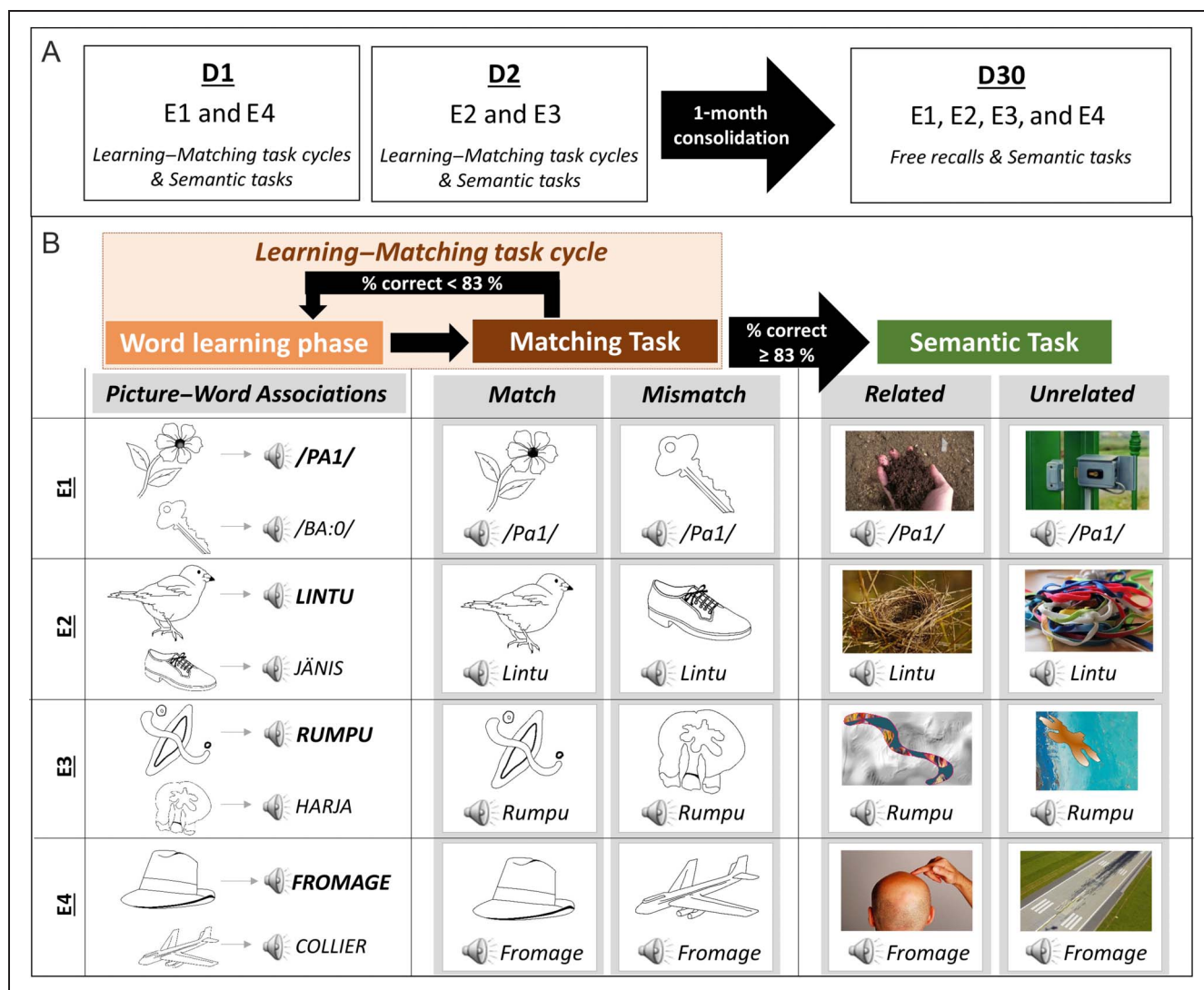


Figure 1. Experimental design. (A) Participants performed four independent novel word learning experiments on 2 days (D1 and D2). Levels of performance in all experiments were retested after 30 days. (B) Each experiment was composed of learning–matching task cycles and a semantic task. The cycles included word learning phases during which each word was paired with its respective picture and matching tasks during which words were presented with one of the pictures, either matching or mismatching the previously learned associations. Once the participant reached 83% correct responses, she or he entered the semantic task where words were presented with novel pictures that were either semantically related or unrelated to the novel words.

E1, E2, E3, and E4 again, after a period of 1 month. The aim was to determine whether musicians would remember the previously learned picture–word associations better than nonmusicians. A positive finding would be taken to reflect enhanced attention top–down processes in musicians compared to nonmusicians and would extend previous results showing enhanced working memory and STM in musicians (Schulze et al., 2012; George & Coch, 2011).

METHODS

Participants

Thirty-three participants contributed data, of which 17 were professional musicians (MUS, eight women) and 16 were nonmusician control participants without formal

music training (NM, eight women), but who were involved in a regular leisure activity (e.g., sports, dance, theater). Of these, two MUS and one NM were considered outliers based on behavioral performance (± 2 SDs away from the mean; see Statistical Analyses section), resulting in equal samples of 15 analyzed data sets for each group. The experiments lasted for altogether 9 hr, and all participants were invited to the laboratory on 3 different days (i.e., two sessions on consecutive days and the third session about 30 days later). The two groups did not significantly differ in age (MUS: mean age = 25.7 years, age range = 19–36 years, $SD = 1.5$; NM: mean age = 26.0 years, age range = 20–35 years, $SD = 1.5$), $F(1, 28) = 0.02$, $p = .90$. All participants were native French speakers, had comparable education levels (university degree), and reported no past or current auditory or neurological deficits. The MUS group practiced

their instruments for an average of 18.3 years (range = 11–29 years, $SD = 5.1$) and included five violinists, three clarinetists, two pianists, two flautists, one oboist, one trombonist, and one trumpeter. None of the participants was bilingual, but all spoke English as a second language. All participants had a basic knowledge of a third language (mainly Spanish, Italian, or German). The study was conducted in accordance with the Declaration of Helsinki concerning human testing. All participants gave their informed written consent before enrolling in the experiment and received monetary compensation for participating.

Screening Measures

Cognitive Ability

Standardized psychometric tests were used to examine verbal STM and working memory (forward and reverse Digit Span, The Wechsler Adult Intelligence Scale [WAIS-III; Wechsler, 1997]), auditory attention (Associated Responses, adapted from the Neuropsychological [NEPSY-II] child battery; Korkman, Kirk, & Kemp, 2007), visual attention (D2-R; Brickenkamp, Schmidt-Atzert, & Liepmann, 2015), lexical and semantic fluency (Verbal Fluency; Cardebat, Doyon, Puel, Goulet, & Joannette, 1990), and nonverbal general intelligence (Matrices, WAIS-III; Wechsler, 1997).

Musical Aptitude

Participants performed two musicality tests (adapted from the Montreal Battery of Evaluation of Amusia (MBEA) battery; Peretz, Champod, & Hyde, 2003) consisting of judging whether pairs of piano melodies were the same or different. One test was based on melodic information, and the other test was based on rhythmic information.

Experimental Stimuli

Auditory Stimuli

For E1, the stimuli were the same as those used by Dittinger et al. (2016). They comprised 12 natural Thai monosyllabic words: /ba1/, /pa1/, /p^ha1/, /ba:1/, /pa:1/, /p^ha:1/, /ba0/, /pa0/, /p^ha0/, /ba:0/, /pa:0/, and /p^ha:0/.¹ These words varied in vowel duration, with short (261 msec, on average) and long (531 msec, on average) vowels, in fundamental frequency, with low-tone ($F_0 = 175$ Hz, on average) and mid-tone ($F_0 = 218$ Hz, on average) vowels, and in voice onset time (/b/ = -144 msec vs. /p/ = 3 msec vs. /p^h/ = 77 msec). Stimuli were recorded by a female Thai–French bilingual, this way ensuring that they were produced naturally. For E2 and E3, 48 natural Finnish disyllabic words were selected: *tyyny, lintu, porsa, kirves, jänis, huntu, pullo, ruoho, sankka, farkut, ryhmä, hanhi, norsu, telta, marsu, hylje, lehmä, laskin, molli, ketju, jalka, myrsky, sänky, huilu, nuoli, lettu, maila, sakset, tähti, harja, kangas, rumpu, juusto, noppa, lisko, sähkö, sydän, varvas, mökki, patja, hillo, kongi, huulet, pyssy, mylly, raastin, järvi, and huivi*. These words included

nonnative phonetic features for French speakers (i.e., geminate stops, e.g., *telta*; initials pronounced as “h,” e.g., *hanhi*; nonnative vowels for French speakers, e.g., *sähkö*) and were presented in the two experiments (i.e., two lists of 24 words counterbalanced across E2 and E3 and across participants). To ensure clear perceptual differences, each list contained words without the same initial syllables. Finally, for E4, 24 natural French disyllabic words without the same initial syllable were chosen: *fléchette, valise, échelle, bougie, balance, collier, girafe, orteil, gilet, abeille, passoire, mouton, soleil, fenêtre, râteau, cactus, pinceau, tortue, maison, asperge, camion, poupée, croissant, and bocal*. A female Finnish–French bilingual recorded the words of E2, E3, and E4, again ensuring that all words were produced naturally. Furthermore, for each word in each experiment, four versions were digitally recorded to reproduce natural speech variability; these different versions of the same word were randomly assigned between word lists and participants. Sound pressure level was normalized across all words to a mean level of 70 dB by using the Praat software (Boersma & Weenink, 2011).

Visual Stimuli

For the learning phases in all four experiments, black and white line drawings were selected. In E1, E2, and E4, these drawings represented familiar objects that were chosen from the Snodgrass and Vanderwart (1980) pictures’ set.² To ensure compatibility with the auditory stimuli, pictures with monosyllabic French names were selected in E1; and pictures with disyllabic French names, in E2 and E4. On the basis of the French normative measures for the Snodgrass and Vanderwart pictures (Alario & Ferrand, 1999), the pictures were controlled for name agreement, image agreement, familiarity, image complexity, age of acquisition, and word frequency. For E1, the same pictures as in Dittinger et al. (2016) were used, whereas for E2 and E4, the remaining pictures were counterbalanced across participants. In E3, line drawings corresponded to the best-rated nonobjects (i.e., do not resemble a real object) created by Kroll and Potter (1984).³ For all four experiments, the same pictures as in the learning phase were then presented in the matching task. For the semantic tasks of E1, E2, and E4, new pictures that the participants had not seen before in the learning–matching task cycles and that were semantically related or unrelated to the meaning of the newly learned words were selected from the Internet. For the semantic task of E3, new visually related or unrelated pictures were created manually in Office PowerPoint (for examples, see Figure 1B). All pictures were pretested with university students ($n = 15$; age range = 18–30 years), who were asked to rate the semantic or visual relatedness between new and old pictures on a scale from 1 to 5 (1 = *not related at all* and 5 = *strongly related*). When the semantic or visual association between the new picture and the old picture was rated as 4 or 5, the pair was later used as a related pair. When the

new–old picture association was rated as 1 or 2, the pair was later used as an unrelated pair.

Experimental Tasks

Participants were tested individually in a quiet and shielded (i.e., Faraday cage) experimental room, where they sat in a comfortable chair at about 1 m from a CRT computer screen. Auditory stimuli were presented through HiFi headphones (Sennheiser, HD590) at 70-dB sound pressure level. Visual and auditory stimuli presentation, as well as the collection of behavioral data, was controlled via the Presentation software (NeuroBehavioral Systems, Version 11.0). Four independent experiments were performed, two on D1 (always E1 and E4, with E1 first) and two on Day 2⁴ (always E2 and E3, with E2 first), and each experiment comprised Learning–Matching task cycles, followed by the semantic task (see Figure 1A and B). Moreover, to test for long-term memory effects, participants performed again the semantic tasks of all four experiments (without learning–matching task cycles) around 30 days (range: 24–41 days) after D1.

Learning–Matching Task Cycles

One learning–matching task cycle consisted of one learning block (i.e., one presentation of each picture–word association), followed by one matching task (i.e., one match and one mismatch condition for each picture–word association). The number of cycles varied across participants based on the percentage of correct responses in the matching task. That is, participants performed learning–matching task cycles until they reached at least 83% correct responses twice in a row (i.e., until participants correctly answered 10 of 12 picture–word associations twice in a row). Then, the semantic task was presented (described below).

Word learning phase. Participants were asked to learn the meaning of the novel words using picture–word associations. In E1, 12 picture–word associations had to be learned, based on known pictures and natural Thai monosyllabic words previously used by Dittinger et al. (2016). In E2, E3, and E4, 24 picture–word associations had to be learned, with known pictures and natural Finnish disyllabic words in E2, nonobjects and natural Finnish disyllabic words in E3, and known pictures and known French disyllabic words in E4 (see Figure 1B for the full list). For instance, in E1, a drawing of a flower was followed by the auditory presentation of the word /pa1/ and, thus, /pa1/ was the word for flower in our “foreign” language. Similarly, in E2, a drawing of a bird was followed by the word “Lintu”; therefore, “Lintu” was the word for bird. In E3, participants were asked to learn the visual appearance (given by the picture of the nonobject) of the novel word “Rumpu.” Finally, in E4, participants learned that “fromage” (meaning cheese in French) was the word for hat. The number of associations to be learned in each experiment (12 in E1 and 24

in E2, E3, and E4) was determined based on pilot data showing that associations were more difficult to learn in E1 than in the other three experiments. In one block of learning, each picture–word association was presented once, resulting in 12 trials for E1 and 24 trials for E2, E3, and E4. The picture was presented first and followed after 750 msec by one of the words. The total trial duration was 2000 msec. Different lists were built, so that across participants, different pictures were associated with different words. No behavioral response was required from the participants during this word learning phase. The duration of one learning block was about half a minute for E1 and 1 min for each of E2, E3, and E4.

Matching task. One of the pictures was presented, followed after 750 msec by an auditory word that either matched or mismatched the association previously learned (in the word learning phase). After having heard the word, participants were asked to press one of two response keys as quickly and accurately as possible, to indicate whether the presented association was correct or incorrect. After 2750 msec, a row of “Xs” appeared on the screen for 1000 msec, and participants were asked to blink during this period, to minimize eye movements during the next trial. The total trial duration was 3750 msec. For instance, in E1, the drawing of a flower followed by /pa1/ (i.e., flower) was a match and the drawing of a key followed by /pa1/ was a mismatch (see Figure 1B for more examples in E2, E3, and E4). The response hands for indicating matching/mismatching responses were counterbalanced across participants. During one block of the matching task, each word was presented twice, once in a match condition and once in a mismatch condition. Thus, 24 trials for E1 and 48 trials for E2, E3, and E4 were presented in one block, whereas the order was pseudorandomized such that no more than four successive “matches” or “mismatches” were presented. The duration of one matching block was 1.5 min in E1 and 3 min in E2, E3, and E4.

Semantic Task

In E1, E2, and E4, new pictures were presented that were semantically related or unrelated to the words used previously in each of these experiments. In E3, new pictures were visually similar or not to the pictures used previously to learn the picture–word associations. In all four experiments, pictures were presented first followed after 1500 msec by the auditory word. Similar to the matching task, after having heard the word, participants were asked to press one of two response keys as quickly and accurately as possible, to indicate whether the picture and the words were related or unrelated. After 3500 msec, a row of “Xs” appeared on the screen for 1000 msec, and participants were asked to blink during this period. The total trial duration was 4500 msec, and the response hands were counterbalanced across participants. For instance, whereas the picture of soil was semantically related to the previously

Table 1. Results of Statistical Analyses on ERRs, RTs, Error Increase, and RT Decrease, in the Different Tasks (Matching, Semantic D1, and Semantic D30 vs. D1) and Experiments (E1, E2, E3, and E4)

		A. Matching Task						B. Semantic Task D1						C. Semantic Task D30 vs. D1					
		ERRs			RTs			ERRs			RTs			Error Increase			RT Decrease		
		<i>F</i> (1, 28)	<i>p</i>	η_p^2	<i>F</i> (1, 28)	<i>p</i>	η_p^2	<i>F</i> (1, 28)	<i>p</i>	η_p^2	<i>F</i> (1, 28)	<i>p</i>	η_p^2	<i>F</i> (1, 28)	<i>p</i>	η_p^2	<i>F</i> (1, 28)	<i>p</i>	η_p^2
Main effect of Group	E1	0.42	.52	.014	0.00	.99	.000	0.32	.58	.011	0.07	.79	.000	0.82	.37	.028	0.10	.75	.003
	E2	0.69	.41	.024	0.51	.48	.017	1.14	.29	.039	0.27	.61	.009	7.50	.01	.211	2.90	.10	.093
	E3	0.02	.88	.000	1.10	.31	.037	0.02	.90	.000	0.01	.94	.000	4.93	.04	.149	1.63	.21	.005
	E4	0.39	.54	.013	1.45	.24	0.49	1.55	.22	.052	3.47	.07	.002	5.56	.03	.165	7.16	.01	.203
Main effect of Condition	E1	5.71	.02	.169	4.21	.05	.130	4.76	.04	.145	49.46	.001	.638	1.18	.29	.040	15.93	.001	.362
	E2	38.45	.001	.579	0.40	.53	.014	68.34	.001	.709	67.26	.001	.706	5.86	.02	.173	23.12	.001	.452
	E3	41.87	.001	.599	0.18	.67	.006	75.01	.001	.728	73.89	.001	.725	6.94	.01	.198	21.17	.001	.430
	E4	25.65	.001	.478	2.90	.10	.093	66.03	.001	.702	41.80	.001	.598	22.57	.001	.446	8.94	.006	.242
Group × Condition	E1	0.24	.63	.008	0.00	.98	.000	0.01	.96	.000	0.03	.86	.001	9.16	.005	.246	0.43	.52	.015
	E2	0.02	.88	.000	1.02	.32	.035	0.80	.38	.027	0.25	.62	.008	3.59	.07	.113	0.40	.53	.014
	E3	0.01	.93	.000	2.83	.10	.091	0.02	.88	.000	0.83	.37	.028	2.39	.13	.094	1.02	.32	.035
	E4	1.30	.26	.044	0.04	.84	.001	0.00	.99	.000	1.51	.23	.051	.27	.60	.009	0.94	.34	.032
Main Effect of First/Final	E1	365.63	.001	.928	139.85	.001	.833												
	E2	337.20	.001	.923	159.91	.001	.850												
	E3	302.00	.001	.915	137.80	.001	.831												
	E4	54.46	.001	.660	97.81	.001	.777												
Group × First/Final	E1	1.81	.19	.060	2.25	.14	.074												
	E2	2.03	.17	.067	0.03	.87	.001												
	E3	1.11	.30	.038	0.14	.71	.004												
	E4	0.69	.41	.024	1.42	.24	.048												

F values, *p* values, and effect sizes (η_p^2) are indicated, and significant *p* values are in **bold**.

learned word /pa1/ (i.e., “flower”) in E1, the picture of a lock was semantically unrelated to /pa1/ (see Figure 1B for more examples in E2, E3, and E4). In all experiments, the semantic task started with a short practice block containing four trials, to familiarize participants with the task. After that, 144 trials grouped in two blocks were presented in E1, and 288 trials grouped in four blocks were presented in each of E2, E3, and E4. In total and for each experiment, every previously learned word was presented 12 times, but none of the new pictures was repeated, so that on each trial, the word was associated with a different related or unrelated picture. Half of the picture–word pairs were semantically or visually related, and half were semantically or visually unrelated. The trial order was pseudorandomized such that no more than four successive “related” or “unrelated” pairs were presented. The duration of one block was 5.4 min.

Long-Term Memory Session

To test for long-term memory of the novel words, participants came back to the laboratory about 1 month after the first experimental session, to perform the semantic tasks from all four experiments again (as described above).

The order of experiments was counterbalanced across participants for this session.

EEG Data Acquisition and Analysis

The EEG was continuously recorded at a sampling rate of 512 Hz with a band-pass filter of 0–102.4 Hz by using a Biosemi amplifier system (BioSemi Active 2) with 32 active Ag/Cl electrodes (Biosemi Pintype), located at standard positions according to the International 10–20 System (Jasper, 1958). The EOG was recorded from flat-type active electrodes placed 1 cm to the left and right of the external canthi and from an electrode placed beneath the right eye. Two additional electrodes were placed on the left and right mastoids. Electrode impedance was kept below 5 k Ω . EEG data were analyzed using the Brain Vision Analyzer software (Version 1.05.0005; Brain Products, GmbH). All data were rereferenced offline to the averaged left and right mastoids and filtered with a bandpass filter from 0.1 to 40 Hz. An independent component analysis and an inverse independent component analysis were used to identify and remove components associated with vertical and horizontal ocular movements. Data were segmented into 1200-msec epochs, time-locked to word onset, and included a 200-msec

Table 2. Matching Tasks: ERRs and RTs (Standard Deviations in Parentheses)

		<i>E1</i>	<i>E2</i>	<i>E3</i>	<i>E4</i>
ERRs (%)	MUS	27.0 (1.2)	21.7 (1.5)	22.1 (1.4)	9.2 (1.7)
	NM	28.1 (1.2)	19.9 (1.5)	22.4 (1.4)	7.7 (1.6)
	First cycles	44.8 (2.1)	34.0 (2.4)	36.3 (2.3)	15.2 (2.9)
	Final cycles	10.3 (1.1)	7.6 (1.0)	8.3 (1.1)	1.6 (0.7)
	Match overall	24.7 (2.1)	27.0 (2.1)	29.1 (1.9)	11.2 (2.2)
	First cycles	43.1 (2.4)	44.0 (2.3)	48.0 (2.3)	20.4 (2.8)
	Final cycles	6.3 (1.2)	10.0 (1.1)	10.1 (1.1)	2.0 (0.6)
	Mismatch overall	30.4 (2.0)	14.6 (2.0)	15.5 (2.2)	5.7 (1.4)
	First cycles	46.5 (1.2)	24.0 (2.5)	24.6 (2.6)	10.1 (1.7)
	Final cycles	14.3 (1.6)	5.1 (0.7)	6.4 (0.9)	1.3 (0.5)
RTs (msec)	MUS	1212 (42)	1130 (42)	1130 (45)	906 (36)
	NM	1212 (42)	1172 (42)	1197 (45)	844 (36)
	First cycles	1423 (60)	1297 (48)	1310 (52)	981 (44)
	Final cycles	1002 (34)	1006 (43)	1017 (46)	770 (34)
	Match overall	1189 (48)	1147 (40)	1166 (47)	862 (35)
	First cycles	1398 (50)	1296 (32)	1313 (40)	973 (29)
	Final cycles	980 (26)	998 (33)	1019 (32)	751 (26)
	Mismatch overall	1235 (41)	1156 (47)	1161 (46)	889 (40)
	First cycles	1448 (43)	1298 (38)	1308 (35)	988 (36)
	Final cycles	1023 (28)	1014 (33)	1015 (34)	790 (25)

baseline. DC-detrend and removal of artifacts above a gradient criterion of 10 $\mu\text{V}/\text{msec}$ or a max–min criterion of 100 μV over the entire epoch were applied. Only trials in which participants gave a correct response were included in the averages, and the number of accepted epochs varied between around 60% (E1) and 80% (E2, E3, and E4). Averages were computed for each participant and for each condition in each experiment, and these individual averages were then averaged into the grand averages across all participants.

Statistical Analyses

ANOVAs were computed using the Statistica software (Version 12.0, StatSoft Inc.). ANOVAs always included Group (MUS vs. NM) as a between-participant factor and specific within-participant factors for each task. Regarding the behavioral data registered on D1, univariate ANOVAs

(only including the Group factor) were computed for all experiments on the number of learning–matching task cycles. For the matching and semantic tasks, ANOVAs were computed on error rates (ERRs) and on RTs and included Condition (Matching task: match vs. mismatch and Semantic task: related vs. unrelated) as a within-participant factor. For the matching task, Cycles (first vs. final cycles⁵) was included as an additional within-participant factor. ANOVAs were first computed by considering each experiment separately and second by including Experiment (E1 vs. E2 vs. E3 vs. E4) as an additional within-participant factor.

Regarding long-term memory for novel words, ANOVAs for D30 always included Group as a between-participant factor and were computed on the difference in ERRs and in RTs between D30 and D1, for both related and unrelated words. As for D1, ANOVAs were first computed separately for each experiment and, second, by including Experiment (E1 vs. E2 vs. E3 vs. E4) as an additional within-participant factor.

Table 3. Semantic Tasks: ERRs and RTs (Standard Deviations in Parentheses) on D1 and D30, and Error Increase and RT Decrease from D1 to D30

		<i>E1</i>	<i>E2</i>	<i>E3</i>	<i>E4</i>
ERRs: D1	MUS	33.5 (1.5)	17.9 (1.8)	24.9 (2.0)	17.2 (2.3)
	NM	34.7 (1.5)	15.1 (1.9)	24.5 (2.0)	13.1 (2.3)
	Related	37.7 (1.7)	25.2 (2.1)	36.6 (2.5)	23.0 (2.4)
	Unrelated	30.4 (2.2)	7.8 (1.0)	12.8 (1.4)	7.3 (1.1)
ERRs: D30	MUS	33.8 (2.3)	26.0 (2.7)	33.3 (2.2)	27.9 (3.2)
	NM	37.1 (2.3)	32.1 (2.7)	39.5 (2.2)	33.6 (3.2)
	Related	41.1 (2.6)	42.1 (3.1)	53.0 (2.4)	47.2 (3.4)
	Unrelated	29.8 (2.7)	16.0 (2.5)	19.8 (2.9)	14.3 (2.7)
Error increase: D1–D30	MUS	0.3 (1.7)	8.1 (2.3)	8.4 (2.1)	10.7 (2.9)
	NM	2.5 (1.7)	17.0 (2.3)	14.9 (2.1)	20.4 (2.9)
	Related	3.4 (2.3)	16.9 (2.9)	16.4 (2.7)	24.2 (3.1)
	Unrelated	−0.6 (2.1)	8.2 (1.9)	7.0 (1.9)	7.0 (2.4)
RTs: D1	MUS	1308 (45)	1352 (53)	1297 (47)	1339 (50)
	NM	1325 (46)	1314 (52)	1292 (46)	1207 (51)
	Related	1242 (31)	1271 (36)	1246 (33)	1232 (34)
	Unrelated	1392 (36)	1395 (40)	1342 (33)	1315 (39)
RTs: D30	MUS	1156 (54)	1269 (54)	1222 (56)	1254 (59)
	NM	1153 (54)	1325 (54)	1277 (56)	1258 (59)
	Related	1121 (39)	1292 (41)	1243 (41)	1249 (45)
	Unrelated	1189 (39)	1302 (38)	1255 (39)	1263 (42)
RT decrease: D1–D30	MUS	−152 (43)	−83 (40)	−75 (33)	−85 (36)
	NM	−172 (44)	12 (40)	−16 (33)	51 (36)
	Related	−121 (34)	21 (30)	−3 (28)	18 (27)
	Unrelated	−203 (30)	−93 (30)	−88 (23)	−52 (29)

Regarding the ERP analyses, for each experiment and on the basis of previous results and visual inspection of the ERP waveforms, the N400 component was analyzed in the semantic task by computing the mean amplitude in the 400- to 550-msec time window. Only correct responses were considered in these analyses. ANOVAs always included Group (MUS vs. NM) as a between-participant factor and Condition (related vs. unrelated) as a within-participant factor, together with Laterality (left: F3, C3, P3; midline: Fz, Cz, Pz; right: F4, C4, P4) and Anterior/Posterior (frontal: F3, Fz, F4; central: C3, Cz, C4; parietal: P3, Pz, P4). As for the behavioral analyses, ANOVAs were first computed for each experiment separately and, second, by including Experiment (E1 vs. E2 vs. E3 vs. E4) as an additional within-participant factor. Post hoc Tukey tests (reducing the probability of Type I errors) were used to determine the origin of significant main effects and interactions. Statistical significance was defined at the .05 alpha level, and results are reported together with the partial eta-squared effect sizes (η_p^2).

RESULTS

Psychometric Measures

Results of univariate ANOVAs showed no significant between-group differences regarding the nonverbal general intelligence ($F < 1$), auditory or visual attention ($F(1, 28) = 1.55, p = .22, \eta_p^2 = .052$, and $F(1, 28) = 1.96, p = .17, \eta_p^2 = .065$, respectively), or lexical or semantic fluency (both $F_s < 1$). By contrast, MUS showed better working memory and STM abilities than NM, $F(1, 28) = 6.90, p = .01, \eta_p^2 = .197$, and $F(1, 28) = 12.83, p = .001, \eta_p^2 = .314$, respectively.

Musicality Task

Results of a 2×2 ANOVA (2 Groups [MUS vs. NM] \times 2 Tasks [Melodic vs. Rhythmic]) showed that MUS made fewer errors (8.1%, $SD = 1.6$) than NM (15.2%, $SD = 1.6$; main effect of Group: $F(1, 28) = 9.32, p = .005, \eta_p^2 = .249$), and all participants performed better on the rhythmic (9.3%, $SD = 1.1$)

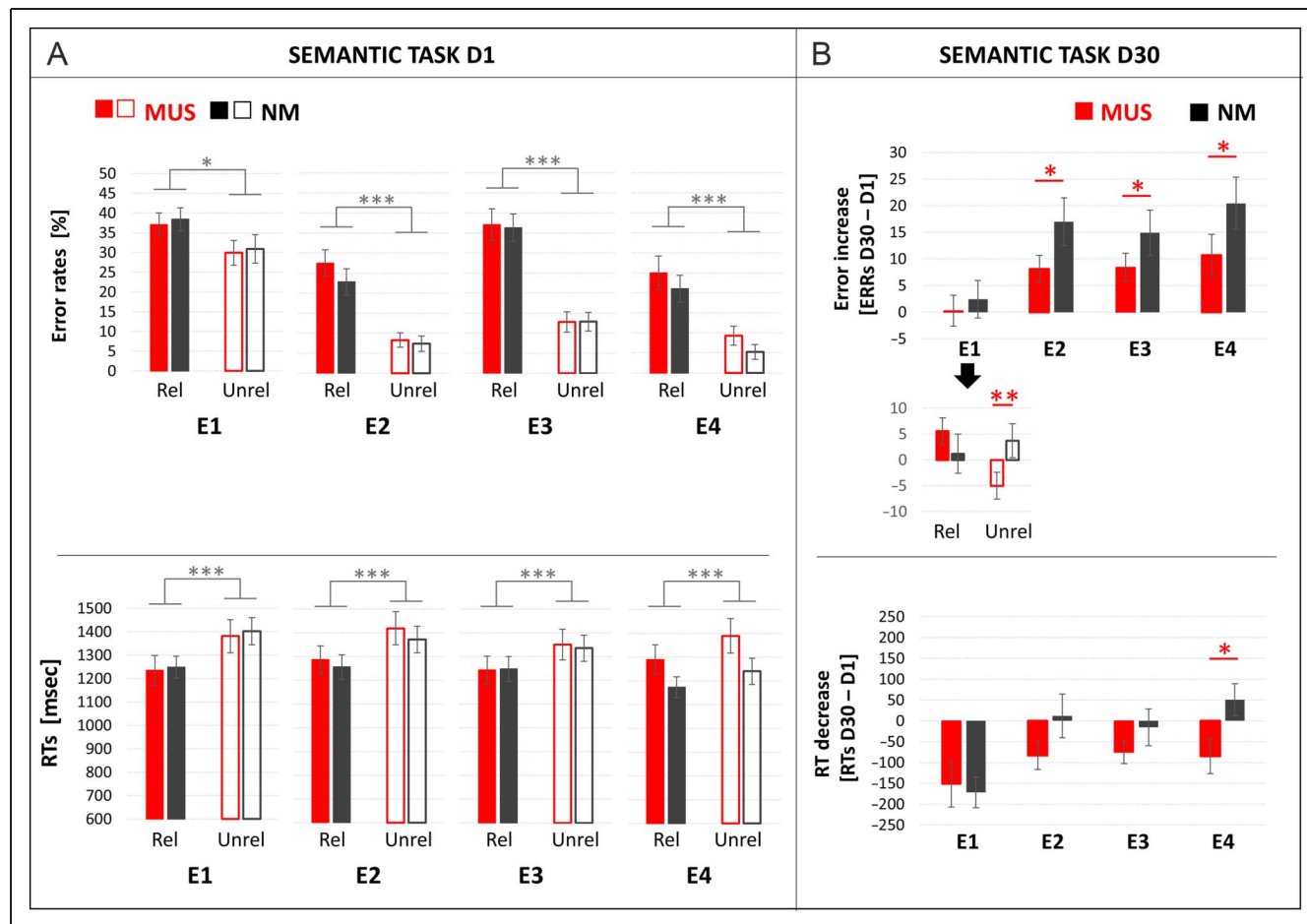


Figure 2. (A) ERRs (top) and RTs (bottom) in the semantic tasks on D1 are shown separately for the four experiments. Results for related (Rel; full bars) and unrelated (Unrel; empty bars) words are illustrated for musicians (MUS; red) and nonmusicians (NM; black). (B) Error increase (top) and RT decrease (bottom) from D1 to D30 in the semantic tasks are shown. The level of significance is represented by asterisks with $*p < .05$, $**p < .01$, and $***p < .001$.

than on the melodic (14.1%, $SD = 1.8$) task (main effect of Task: $F(1, 28) = 6.56, p = .02, \eta_p^2 = .189$). No Group \times Task interaction was observed.

Learning–Matching Task Cycles: Behavioral Results

Number of Cycles

Results of univariate ANOVAs (including Group: MUS vs. NM) for each experiment showed that the number of cycles necessary to reach the threshold did not significantly differ for MUS and for NM in any of the four experiments (main effects of Group: E1, E2, E3, E4: all $F_s < 1$). Moreover, the between-experiments comparison using two-way ANOVAs (2 Groups \times 4 Experiments) showed that participants needed more cycles in E1 (9.0 cycles, $SD = 0.5$) and fewer cycles in E4 (2.4 cycles, $SD = 0.1$), compared to E2 (3.6 cycles, $SD = 0.2$; Tukey, $p < .001$ and $p = .02$, respectively) and E3 (3.8 cycles, $SD = 0.2$; Tukey, $p < .001$ and $p = .005$, respectively), which did not differ from each

other (number of cycles: $E1 > E2 = E3 > E4$; main effect of Experiment: $F(3, 84) = 105.08, p < .001, \eta_p^2 = .789$).

Matching Task

Results of separate ANOVAs (2 Groups [MUS vs. NM] \times 2 Cycles [First vs. Final] \times 2 Conditions [Match vs. Mismatch]; see Table 1A for F values, p values, and effect sizes [η_p^2] and Table 2 for mean ERR and RT values) for each experiment showed that ERRs and RTs did not significantly differ between MUS and NM. Moreover, in all four experiments, participants made fewer errors and responded faster in the final compared to the first cycles.

Regarding the comparison between experiments (2 Groups [MUS vs. NM] \times 4 Experiments [E1 vs. E2 vs. E3 vs. E4] \times 2 Conditions [Match vs. Mismatch] \times 2 Cycles [First vs. Final]) in the first cycles, ERRs were highest and RTs were slowest in E1, intermediate in E2 and E3, and lowest and fastest in E4 (matching task: $E1 > E2 = E3 > E4$;

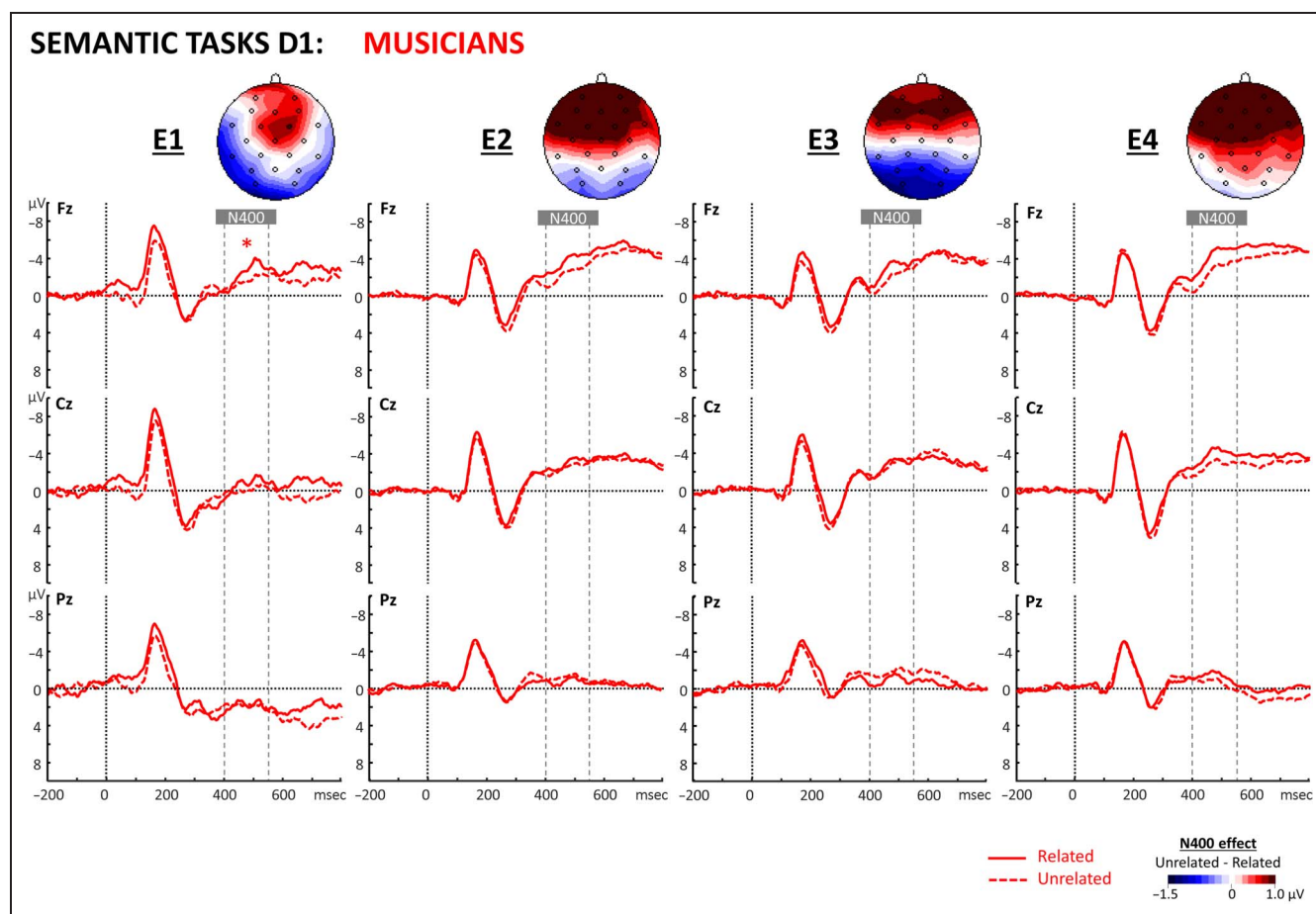


Figure 3. Semantic tasks of the four experiments (E1, E2, E3, and E4) on D1 are shown for musicians. ERPs recorded at frontal (Fz), central (Cz), and parietal (Pz) sites are overlapped for semantically related (solid-red lines) and unrelated (dashed-red lines) words. In this and subsequent figures, time in milliseconds is in abscissa and the amplitude of the effects in microvolt is in ordinate. Time 0 corresponds to word onset, and negativity is plotted upward. Latency windows for statistical analyses are indicated with gray dotted lines, and the level of significance is represented by asterisks with $*p < .05$, $**p < .01$, and $***p < .001$. Topographic voltage distribution maps of the unrelated minus related differences are illustrated for N400 components. Voltage values are scaled from -1.5 to $+1.0$ μV .

Tukey, ERRs and RTs: all $ps < .001$). By contrast, in the final cycles, ERRs and RTs were not significantly different in E1, E2, and E3 but were still lower and faster in E4 ($E1 = E2 = E3 > E4$; Tukey, ERRs: all $ps < .01$; and RTs: all $ps < .001$; Cycle \times Experiment interactions: ERRs: $F(3, 84) = 29.21, p < .001, \eta_p^2 = .510$; and RTs: $F(3, 84) = 10.87, p < .001, \eta_p^2 = .279$; level of performance: $E4 > E2 = E3 = E1$).

Semantic Task, D1: Behavioral Results

The results of separate ANOVAs (2 Groups [MUS vs. NM] \times 2 Conditions [Related vs. Unrelated]; see Table 1B for F values, p values, and effect sizes [η_p^2]; Table 3 for mean ERR and RT values; and Figure 2A) for each experiment showed that ERRs and RTs did not significantly differ between MUS and NM. Moreover, participants made more errors and responded faster to related than unrelated words in all four experiments.

Regarding the comparison between experiments (i.e., 2 Groups [MUS vs. NM] \times 4 Experiments [E1 vs. E2 vs. E3 vs.

E4] \times 2 Conditions [Related vs. Unrelated]), participants made most errors in E1, intermediate errors in E3, and fewest errors in E2 and E4 (ERRs: $E1 > E3 > E2 = E4$; Tukey, all $ps < .001$; main effect of Experiment: $F(3, 84) = 80.36, p < .001, \eta_p^2 = .741$). RTs were similar in all four experiments for related words (Tukey, all $ps > .10$) but faster in E3 and E4 compared to E1 and E2 for unrelated words (Tukey, all $ps < .001$; Experiment \times Condition: $F(3, 84) = 5.24, p = .002, \eta_p^2 = .157$).

Semantic Task, D1: ERP Results

Turning to the electrophysiological data (see Figure 3 for MUS and Figure 4 for NM), results of the separate ANOVAs for each experiment (2 Groups [MUS vs. NM] \times 2 Conditions [Related vs. Unrelated] \times 3 Laterality [Left vs. Central vs. Right] \times 3 Anterior/Posterior [Frontal vs. Central vs. Parietal]; see Table 4 for μV values and Figures 3 and 4) showed that, in E1, the Group \times Anterior/Posterior and Group \times Condition \times Laterality interactions were significant, $F(2,$

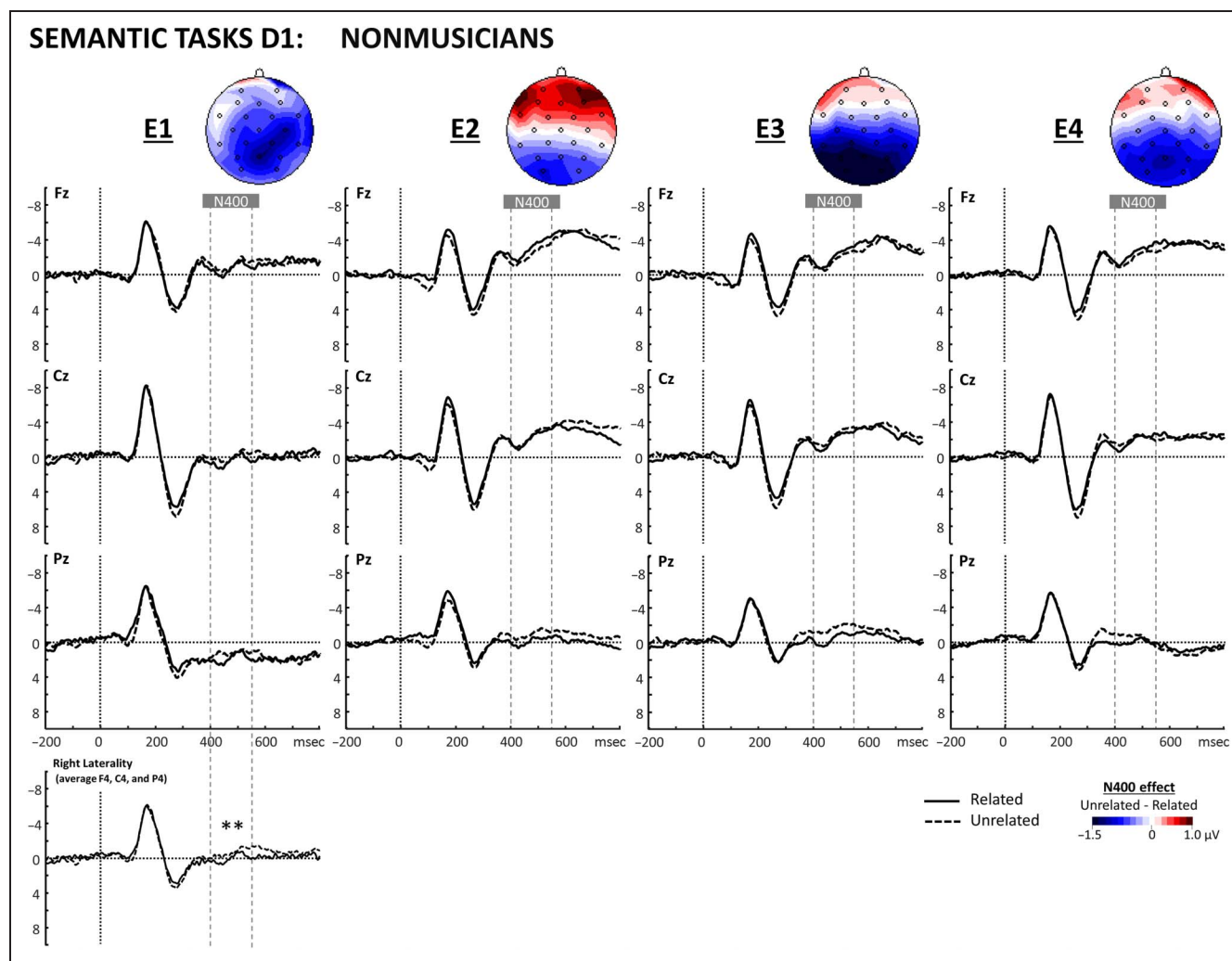


Figure 4. Semantic tasks of the four experiments (E1, E2, E3, and E4) on D1 are shown for nonmusicians. ERPs recorded at frontal (Fz), central (Cz), and parietal (Pz) sites are overlapped for semantically related (solid-black lines) and unrelated (dashed-black lines) words. For E1, an additional overlap of related and unrelated words is shown over the right hemisphere (average of F4, C4, and P4).

Table 4. Semantic Tasks on D1 and D30: N400 Amplitudes in μV (Standard Deviations in Parentheses)

		<i>E1</i>	<i>E2</i>	<i>E3</i>	<i>E4</i>
D1	Related	-0.24 (1.50)	-2.17 (1.28)	-1.93 (1.08)	-2.33 (1.56)
	Frontal	-1.76 (0.97)	-3.38 (0.82)	-2.66 (0.75)	-3.46 (0.97)
	Central	-0.46 (0.93)	-2.58 (0.81)	-2.35 (0.66)	-2.80 (0.97)
	Parietal	1.50 (0.83)	-0.55 (0.74)	-0.78 (0.60)	-0.74 (0.89)
	Unrelated	-0.32 (1.55)	-2.00 (1.16)	-2.11 (1.09)	-1.80 (1.48)
D30	Related	0.69 (1.50)	-0.79 (1.26)	-0.68 (1.30)	-0.62 (1.24)
	Frontal	-0.52 (0.96)	-1.55 (0.83)	-1.43 (0.86)	-1.43 (0.82)
	Central	0.34 (0.93)	-1.16 (0.81)	-1.10 (0.76)	-0.96 (0.78)
	Parietal	2.26 (0.87)	0.34 (0.71)	0.51 (0.75)	0.52 (0.73)
	Unrelated	0.57 (1.44)	-1.03 (1.25)	-1.68 (1.09)	-1.33 (1.24)
	Frontal	-0.35 (0.85)	-1.49 (0.76)	-1.82 (0.76)	-1.93 (0.87)
	Central	0.33 (0.88)	-1.38 (0.79)	-2.06 (0.66)	-1.79 (0.78)
	Parietal	1.73 (0.89)	-0.22 (0.73)	-1.15 (0.61)	-0.29 (0.67)

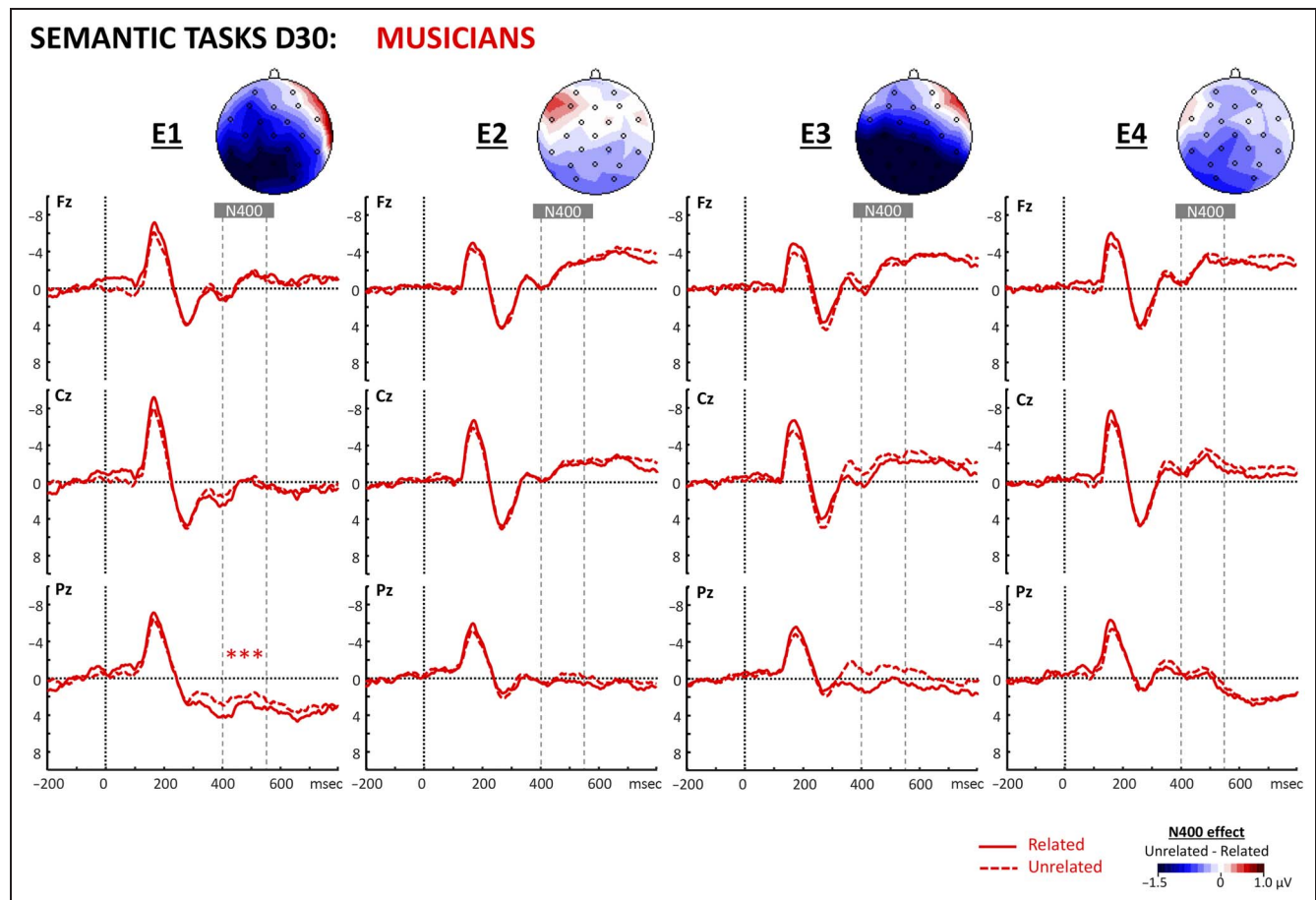


Figure 5. Semantic tasks of the four experiments (E1, E2, E3, and E4) on D30 are shown for musicians. ERPs recorded at frontal (Fz), central (Cz), and parietal (Pz) sites are overlapped for semantically related (solid-red lines) and unrelated (dashed-red lines) words.

56) = 3.40, $p = .04$, $\eta_p^2 = .108$, and $F(2, 56) = 3.23$, $p = .05$, $\eta_p^2 = .103$, respectively. Separate ANOVAs revealed larger N400 to related than unrelated words over frontal sites in MUS (related: $-2.62 \mu\text{V}$, $SD = 1.43$; unrelated: $-1.90 \mu\text{V}$, $SD = 1.40$; Tukey, $p = .02$; Condition \times Anterior/Posterior: $F(2, 28) = 5.19$, $p = .01$, $\eta_p^2 = .270$). By contrast, in NM, N400 was larger to unrelated than related words over mid-line and right hemisphere (unrelated: $-0.03 \mu\text{V}$, $SD = 1.37$, and $-0.78 \mu\text{V}$, $SD = 1.16$, respectively; related: $0.53 \mu\text{V}$, $SD = 1.45$, and $-0.01 \mu\text{V}$, $SD = 1.20$, respectively; Tukey, $p = .03$ and $p = .002$, respectively; Condition \times Laterality: $F(2, 28) = 3.32$, $p = .05$, $\eta_p^2 = .191$; see Figure 4, "E1"). In E2, E3, and E4, neither the main effect of Group nor any interactions involving the Group factor were significant. Finally, in E1, E2, and E3, the N400 was larger to unrelated than related words over parietal sites and to related than unrelated words over frontal sites in E2 and E3 (Condition \times Anterior/Posterior interactions, E1: $F(2, 56) = 3.27$, $p = .05$, $\eta_p^2 = .104$; E2: $F(2, 56) = 44.93$, $p < .001$, $\eta_p^2 = .616$; and E3: $F(2, 56) = 53.09$, $p < .001$, $\eta_p^2 = .654$; Tukey, E1: $p = .01$, E2 and E3: all $ps < .001$). In E4, the N400 was larger to related than unrelated words over frontal and central sites (Tukey, both $ps < .001$; Condition \times Anterior/Posterior: $F(2, 56) = 21.62$, $p < .001$, $\eta_p^2 = .435$; see Figure 7, "D1").

Regarding the comparison between experiments (2 Groups [MUS vs. NM] \times 4 Experiments [E1 vs. E2 vs. E3 vs. E4] \times 2 Conditions [Related vs. Unrelated] \times 3 Laterality [Left vs. Central vs. Right] \times 3 Anterior/Posterior [Frontal vs. Central vs. Parietal]), N400 amplitude was smaller in E1 than in E2, E3, and E4 (Tukey, all $ps < .001$; main effect of Experiment: $F(3, 84) = 16.91$, $p < .001$, $\eta_p^2 = .376$).

Semantic Task: Change in Behavioral Performance from D1 to D30

The results of separate ANOVAs for each experiment (2 Groups [MUS vs. NM] \times 2 Conditions [Related vs. Unrelated]; see Table 1C for F values, p values, and effect sizes [η_p^2]; Table 3 for mean ERR and RT values; and Figure 2B) revealed that, in E2, E3, and E4, the error increase from D1 to D30 was smaller in MUS than in NM (main effects of Group, E2: $F(1, 28) = 7.50$, $p = .01$, $\eta_p^2 = .211$; E3: $F(1, 28) = 4.93$, $p < .04$, $\eta_p^2 = .149$; and E4: $F(1, 28) = 5.56$, $p < .03$, $\eta_p^2 = .165$). In E1, the group effect was significant for unrelated words only (Group \times Condition interaction: $F(1, 28) = 9.16$, $p = .005$, $\eta_p^2 = .246$, with a significant main effect of Group for unrelated words, $F(1, 28) = 8.98$, $p = .006$, $\eta_p^2 = .242$, but not for related words, $F(1, 28) =$

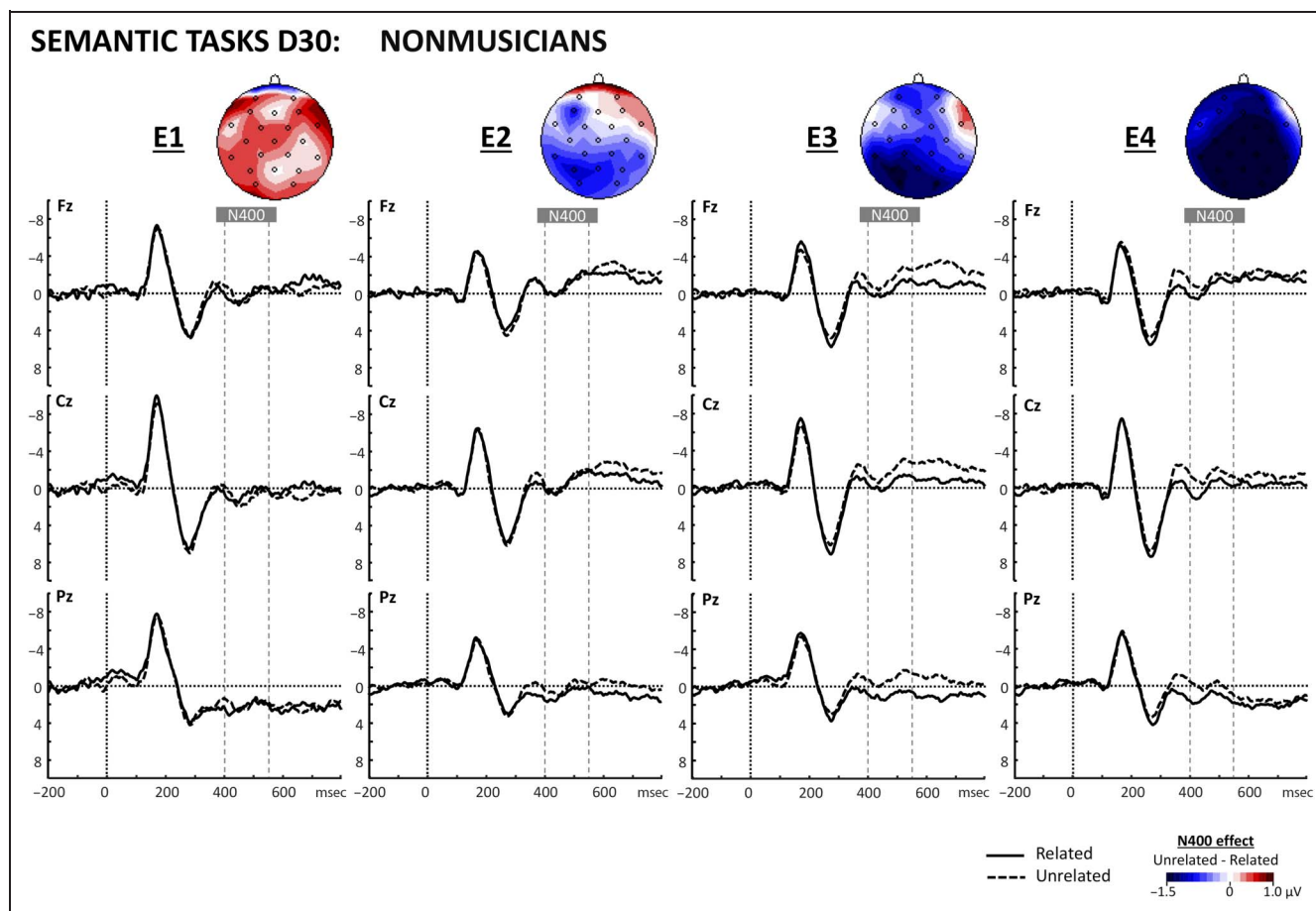


Figure 6. Semantic tasks of the four experiments (E1, E2, E3, and E4) on D30 are shown for nonmusicians. ERPs recorded at frontal (Fz), central (Cz), and parietal (Pz) sites are overlapped for semantically related (solid-black lines) and unrelated (dashed-black lines) words.

3.84, $p = .06$, $\eta_p^2 = .120$). RT decreases from D1 to D30 were larger in MUS than in NM only in E4 (main effect of Group: $F(1, 28) = 7.16$, $p < .01$, $\eta_p^2 = .203$).

Regarding the comparison between experiments (2 Groups [MUS vs. NM] \times 4 Experiments [E1 vs. E2 vs. E3 vs. E4] \times 2 Conditions [Related vs. Unrelated]), error increases were smaller and RT decreases were larger in E1 compared to the other three experiments (Tukey, all $ps < .001$). Moreover, error increases were smaller in E3 compared to E4 (Tukey, $p = .05$; main effects of Experiment: ERRs: $F(3, 84) = 33.90$, $p < .001$, $\eta_p^2 = .547$; and RTs: $F(3, 84) = 9.51$, $p < .001$, $\eta_p^2 = .253$).

Semantic Task, D30: ERP Results

The results of separate ANOVAs for each experiment (2 Groups [MUS vs. NM] \times 2 Conditions [Related vs. Unrelated] \times 3 Laterality [Left vs. Central vs. Right] \times 3 Anterior/Posterior [Frontal vs. Central vs. Parietal]; see Table 4 for μV values, Figure 5 for MUS, and Figure 6 for NM) revealed that the Group \times Condition \times Anterior/Posterior interaction was

significant in E1, $F(2, 56) = 3.27$, $p = .05$, $\eta_p^2 = .104$. Separate ANOVAs showed larger N400 to unrelated than related words over parietal sites in MUS (unrelated: 1.56 μV , $SD = 1.37$; related: 2.56 μV , $SD = 1.45$; Tukey, $p < .001$; Condition \times Anterior/Posterior: $F(2, 28) = 9.86$, $p < .001$, $\eta_p^2 = .413$), but not in NM (main effect of Condition and Condition \times Anterior/Posterior interaction: both $Fs < 1$). In E2, E3, and E4, neither the main effect of Group nor any interaction involving the Group factor was significant. Finally, in E1 and E2, results showed larger N400 to unrelated than related words over parietal sites (Tukey, $p = .009$ and $p = .004$, respectively; Condition \times Anterior/Posterior interactions: $F(2, 56) = 6.13$, $p = .004$, $\eta_p^2 = .179$, and $F(2, 56) = 4.50$, $p = .02$, $\eta_p^2 = .138$, respectively) and over all scalp sites in E3 and E4 (main effect of Condition: $F(1, 28) = 13.53$, $p < .001$, $\eta_p^2 = .325$, and $F(1, 28) = 7.54$, $p = .01$, $\eta_p^2 = .212$, respectively; see Figure 7).

Regarding the comparison between experiments (2 Groups [MUS vs. NM] \times 4 Experiments [E1 vs. E2 vs. E3 vs. E4] \times 2 Conditions [Related vs. Unrelated] \times 3 Laterality [Left vs. Central vs. Right] \times 3 Anterior/Posterior [Frontal vs.

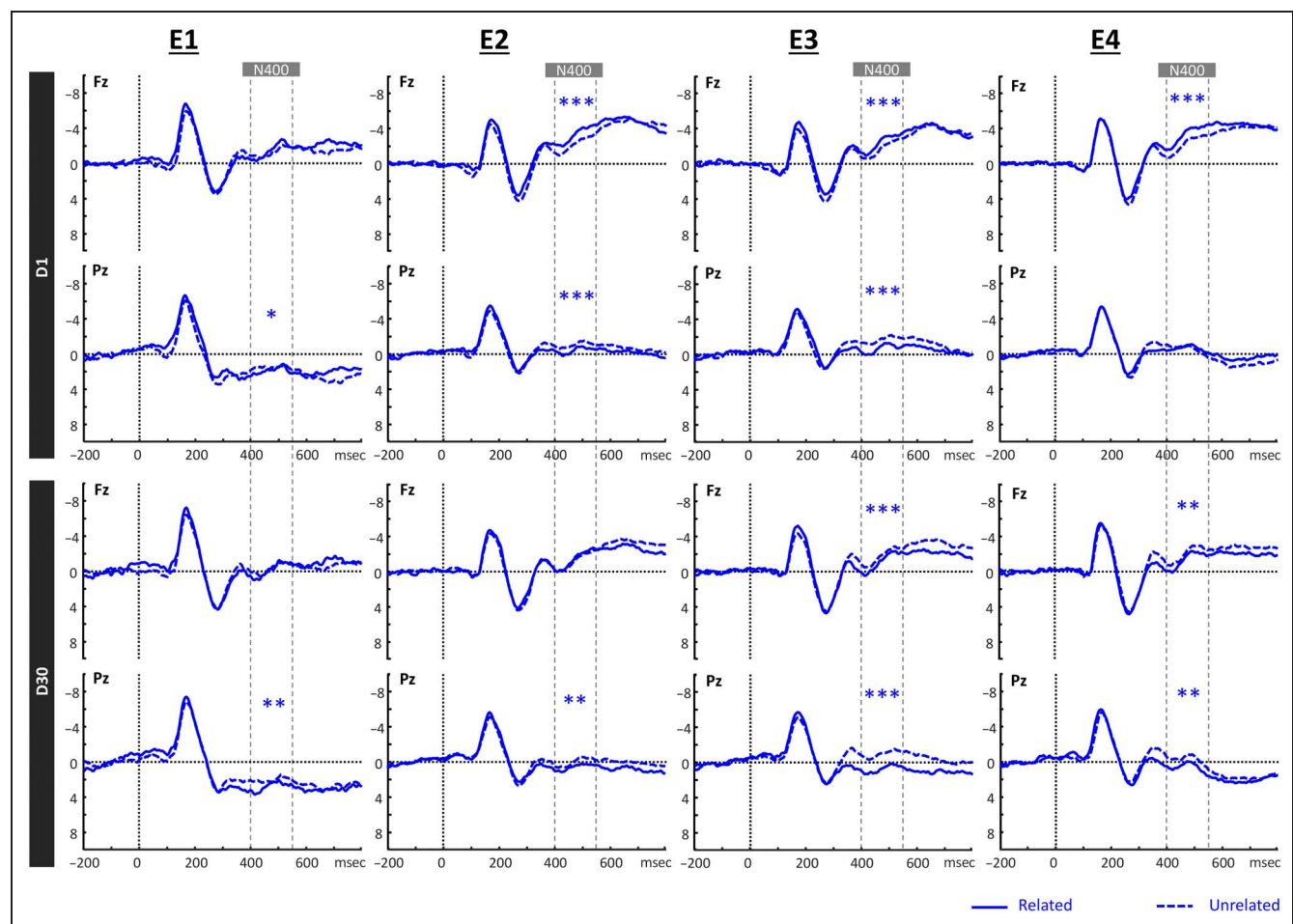


Figure 7. Semantic tasks of the four experiments (E1, E2, E3, and E4) on D1 (top row) and D30 (bottom row) are shown for all participants (averaged across musicians and nonmusicians). ERPs recorded at frontal (Fz) and parietal (Pz) sites are overlapped for semantically related (solid-blue lines) and unrelated (dashed-blue lines) words.

Central vs. Parietal]), the N400s were smaller in E1 than in E2, E3, and E4 (Tukey, all p s < .001; main effect of Experiment: $F(3, 84) = 12.50, p < .001, \eta_p^2 = .308$).

DISCUSSION

On the basis of previous results showing better novel word learning in adult musicians than in nonmusicians (Ditinger et al., 2016), we conducted a series of experiments in which we reduced the influence of auditory perception and associative learning on semantic processing. Of interest was to determine whether, under such conditions, musicians would still outperform nonmusicians in the semantic task, whether audiovisual integration contributed to the musician's advantage in novel word learning, and whether musicians would remember the newly learned words better than nonmusicians 1 month later. Overall, the results of both behavioral and electrophysiological data did not bring evidence for a musician's advantage in semantic processing on D1, thereby showing that, when between-group differences in auditory perception are neutralized, musicians no longer outperform nonmusicians in the semantic task. Interestingly, however, on D30, musicians showed better long-term memory for the novel words. Finally, the results did not reveal better audiovisual integration in musicians than in nonmusicians. These findings are discussed in detail below.

D1: Learning–Matching Task Cycles

To neutralize the influence of auditory perception and associative learning, participants performed a variable number of short learning–matching task cycles until each participant reached a level of 83% correct responses in the matching task. Successful learning in all four experiments was reflected by lower ERRs and faster RTs in the final compared to the first cycles of the matching tasks. Participants needed five cycles on average to reach the learning threshold (i.e., five repetitions of each picture–word association in the learning phases and five repetitions in the matching tasks). Such fast mapping (Carey, 1978) is not surprising in view of the word learning literature evidencing that novel word encoding can be successful only with few repetitions (Borovsky, Elman, & Kutas, 2012; Batterink & Neville, 2011; Dobel et al., 2010; Mestres-Missé, Rodriguez-Fornells, & Münte, 2007; Perfetti, Wlotko, & Hart, 2005).

Interestingly, the number of cycles varied as a function of the experiment, with significantly more cycles in E1 ($n = 9$), intermediate number in E2 and E3 ($n = 3.6$ and $n = 3.8$, respectively), and fewer cycles in E4 ($n = 2.4$; $E1 > E2 = E3 > E4$). Because the same experimental design was used in the four experiments, these differences likely result from the different levels of familiarity and discriminability of the stimuli. Whereas in E4, the novel words were familiar French words, in E1, the novel words were unfamiliar monosyllabic Thai words (including aspiration, tonal, and duration contrasts) that were difficult to discriminate for French

participants, and in E2 and E3, the stimuli were disyllabic Finnish words that were unfamiliar but still easy to discriminate. Thus, participants needed more cycles in E1, even if the number of to-be-learned words was half the number of words in the other three Experiments (i.e., 12 words compared to 24 words in E2, E3, and E4).⁶ Note that results in E1, showing that 18 repetitions of each picture–word association (nine cycles, 18 repetitions) were needed on average to reach the threshold, are comparable to the results reported by Dittinger et al. (2016), in which 20 repetitions of each picture–word association were used to reach a similar level of performance in the matching task. Taken together, these results point to the importance of the stimulus material in driving novel word learning.

The interpretation of E1 being the most difficult and E4 being the easiest experiment is supported by the results in the first cycles of the matching tasks showing the highest ERRs and slowest RTs in E1 as well as the lowest ERRs and fastest RTs in E4 ($E1 > E2 = E3 > E4$). Moreover, whereas E4 remained the easiest task in the final cycles, the level of performance in E1 was similar to that in E2 and E3 (no significant differences on ERRs and on RTs; $E1 = E2 = E3 > E4$), thereby showing that, although the difficult Thai words in E1 needed more repetitions, they were comparably well encoded in the final cycles, relative to the (easier) Finnish words.

Most importantly for the specific goal of this study, ERRs and RTs in the matching task were not significantly different for musicians and nonmusicians in any of the four experiments (see Figure 2A). This is evidence that the implemented manipulation was successful and that potential group differences based on auditory perception, and perceptual learning-related processes (as shown, for instance, by the higher level of performance of musicians compared to nonmusicians in both the rhythmic and melodic tasks) were neutralized before participants performed the semantic task. Nevertheless, the possibility remains that musicians and nonmusicians used different strategies to reach similar outcomes.

D1: Semantic Task

When the perceptual differences between musicians and nonmusicians were neutralized, the results showed no between-group difference at a behavioral level and only in E1 at the electrophysiological level. However, as the E1 results were not expected based on previous results (Dittinger et al., 2016, 2017), they would need to be replicated before being considered further.

In summary, answering our first research question, it is likely that the between-group differences reported in the Dittinger et al. (2016, 2017) studies were driven by better auditory perception (bottom–up processes) in the musicians' group.

Semantic Priming Effects

Typical semantic priming effects (Meyer & Schvaneveldt, 1971) were found in both groups of participants and in all

four experiments, with shorter RTs for related than for unrelated trials (see Figure 2A). This is taken as evidence that novel words were already integrated into semantic networks, so that both musicians and nonmusicians were able to rapidly generalize the novel words' meanings to novel concepts. Moreover, this finding also clearly points to a relationship between word discriminability at the phonological level and word learning at the semantic level. However, in all four experiments, and as reported in the group of young adults tested by Dittinger et al. (2016), participants made more errors to related than unrelated words, thereby producing a speed–accuracy trade-off. This possibly reflects a response bias toward rejection (i.e., considering the word as unrelated to the picture), when the task generates a high degree of response uncertainty (Gigerenzer, 1991).

The interpretation that the novel words were integrated into semantic networks and produced semantic priming effects is supported by the electrophysiological data, showing larger N400 components to the unrelated than to the related words over parietal sites in E1, E2, and E3 and larger N400s to the related than to the unrelated words over frontal sites in E2, E3, and E4 (see Figure 7). In line with previous literature (François et al., 2017; Dittinger et al., 2016; Borovsky et al., 2012; Cunillera et al., 2009; De Diego Balaguer, Toro, Rodriguez-Fornells, & Bachoud-Lévi, 2007; Mestres-Missé et al., 2007), the parietal N400 is taken as evidence that the meaning of the novel words has been integrated into semantic networks, whereas the presence of a frontal reversed N400 is taken to reflect manipulation of novel information in working memory (Hagoort, 2014) and/or ongoing consolidation of episodic memory traces (Mestres-Missé et al., 2007).

Interestingly in E4, results showed larger N400 to related than unrelated words (inversed N400 effects) over fronto-central sites, together with no typical N400 effect over parietal sites. In this experiment (E4), the novel word learning required participants to replace an existing association stored in long-term memory (e.g., the picture of a hat associated with the word hat, “chapeau”) by a new association (e.g., expect the word “fromage,” cheese, after the picture of a hat; see Figure 1). The fronto-central inversed N400 effect is in line with the interpretations proposed above. However, the absence of a parietal N400 effect possibly indicates that the new association (“hat” [picture] – “cheese” [word]) has not yet overridden the old association (“hat” [picture] – “hat” [word]) in long-term memory, even if E4 seemed the easiest of the four experiments performed on D1 (as reflected by fastest learning and lowest ERRs in the matching and semantic tasks).

Finally, we turn to the second question that this series of experiments was designed to answer. On the basis of previous results showing better audiovisual binding (Bidelman, 2016) and higher detection rate of audiovisual violations of a preestablished rule (i.e., “the higher the pitch of the tone, the higher the position of the circle”; Paraskevopoulos et al., 2012) in musicians than in nonmusicians, we asked whether musical expertise would have an influence on

the audiovisual integration when picture–word associations did or did not contain semantic information (control experiments E3 and E4). Results revealed that the musicians did not show improved audiovisual integration, by comparison to the nonmusicians, when the musicians' advantage on auditory perception and associative learning has been neutralized. Thus, coming to an answer, these results suggest that the more efficient audiovisual integration and better novel word learning through picture–word associations reported in musicians than in nonmusicians in previous experiments likely relied on better auditory perception, that is, on bottom–up rather than on top–down process.

D30: Semantic Task

Participants were retested 1 month later, to investigate the long-term memory effects of novel word learning and whether long-term memory for novel words would differ between musicians and nonmusicians. Results showed that musicians outperformed nonmusicians in all four experiments on D30 (see Figure 2A). This was reflected by smaller error increases from D1 to D30 for musicians than for nonmusicians, for both related and unrelated words in all experiments, except in E1, where this effect was only significant for unrelated words. RT decreases from D1 to D30 were also larger in musicians than in nonmusicians in E4. In the ERPs, results in E1 showed between-group differences with larger N400 components to unrelated than related words over parietal sites in musicians, but not in nonmusicians (compare Figures 5 and 6). Thus, in addition to better behavioral performance in all four experiments, the between-group differences on semantic processing on D30 were also reflected by the N400 component, but only when the words were unfamiliar and difficult to discriminate (i.e., Thai words in E1; see discussion above).

At least three interpretations possibly account for better long-term memory in musicians. First, although here there was no evidence showing that musicians encoded the words better at a behavioral level (i.e., no between-group differences in the number of learning cycles, no between-group difference in the matching tasks and in the semantic tasks on D1), results of psychometric tests revealed better working memory and STM abilities in musicians than in nonmusicians, which possibly contributed to form stronger memory traces of the novel words, for later recall (Wojcik, 2013; Morris, Bransford, & Franks, 1977). Thus, words remembered on D30 were possibly more strongly encoded and better integrated into the semantic networks, already on D1. Second, the musicians' higher level of performance on D30 was accompanied by significant N400 effects only in E1. This suggests that the superior auditory skills of musicians may still play a role under difficult encoding conditions of Thai words that called upon preexisting representations of duration and pitch. Thus, in this case, the stronger memory traces of the novel words for musicians than for non-musicians may emerge from better auditory processing, and better auditory processing may also foster recognition of

the Thai stimuli on D30. Finally, it is also possible that long-term consolidation processes are more efficient in musicians than in nonmusicians. Previous results have clearly highlighted the importance of consolidation periods in novel word learning (Bakker, Takashima, van Hell, Janzen, & McQueen, 2015; Takashima, Bakker, van Hell, Janzen, & McQueen, 2014; Dumay & Gaskell, 2007). These different interpretations can be disentangled in further experiments by sorting ERPs recorded during novel word encoding (i.e., learning phase on D1) as a function of whether the words were subsequently remembered on D30 or not (i.e., examining differences at encoding based on subsequent memory, the so-called Dm effect; Paller, Kutas, & Mayes, 1987).

Semantic Priming Effects

The N400 was larger to unrelated than related words across all scalp sites in E3 and E4 and only over parietal sites in E1 and E2 (see Figure 7). These findings suggest that, under ecologically valid conditions for novel word learning and semantic processing (i.e., in E1 and E2, in which novel Thai or Finnish words are attached to known concepts [i.e., meaningful pictures], as it is the case in foreign language learning), N400 effects are localized over parietal scalp sites, as previously reported in the literature (Kutas & Federmeier, 2011; Kutas & Hillyard, 1980). By contrast, results of the control conditions for associative learning (i.e., E3: no semantic information in the novel Finnish words and in the novel nonmeaningful pictures; E4: violations of already well-established semantic associations between French words and meaningful but unrelated pictures) showed largely distributed effects over the scalp, possibly reflecting the call into play of more general processes than the specific semantic processes involved in E1 and E2.

Between-Experiment Comparison

As discussed above, although E1 was very difficult for all participants, the error increase from D1 to D30 was lower, the RT decrease was larger, and the N400 amplitude was smaller than in the other three experiments (see Figures 2B and 7). These results may be explained by the higher number of repetitions in the learning–matching task cycles in E1 than in the other experiments (i.e., nine cycles in E1 vs. 2.4 in E4, 3.6 in E2, and 3.8 in E3). This larger number of repetitions possibly favored deeper encoding and stronger traces in long-term memory. Such an interpretation would also account for the results in E4: Although participants made the fewest errors in E4 compared to the other experiments on D1, the error increase in E4 was higher than in E3 on D30, suggesting that fast learning, with the lowest number of repetitions on D1, was not sufficient to establish strong memory traces. This is an interesting issue for future experiments that may test for the influence of the number of repetitions on long-term memory for words, for example, by testing whether

participants who learn most slowly (i.e., with the highest number of learning–matching task cycles) remember best after 30 days.

Conclusion

By using an experimental design similar to Dittinger et al. (2016), and by comparing four sets of novel words in different experiments, results showed three main findings. First, enhanced auditory perception seem to largely contribute to explain the musicians' advantage in novel word learning. If such advantages are neutralized by testing participants when they have reached the same level of performance in the matching task, results showed no between-group differences in the semantic task. Second, results on D30 clearly showed that fast mapping between a novel word and a concept (represented by a picture) through few repetitions is sufficient to establish a memory trace of the novel word meaning that can be accessed 1 month later. Third and most importantly, results revealed better long-term memory for novel words in musicians than in nonmusicians in all four experiments, a result that, to the best of our knowledge, has not been reported before. These results extend to long-term memory, transfer effects previously described from music expertise to working memory and STM (Kraus, Strait, & Parbery-Clark, 2012; Schulze et al., 2012; George & Coch, 2011; Ho, Cheung, & Chan, 2003; for a meta-analysis, see Talamini, Altoè, Carretti, & Grassi, 2017).

Turning to limitations, the present experiment is based on a correlational design and results do not demonstrate causality. Longitudinal experiments are needed to further test for a causal link between musical expertise and long-term memory. Another limitation is that, because of the complexity of the experimental design involving four experiments and several dependent variables, we only analyzed the N400 and no other ERP components that may have revealed complementary information.

Importantly, whereas results on D1 suggest that the musicians' advantage in novel word learning was mainly driven by auditory perception, results of D30 nevertheless point to differences in long-term memory. Taken together, these findings suggest that enhanced auditory perception, reported in previous experiments (Strait et al., 2010; Micheyl et al., 2006; Kishon-Rabin et al., 2001; Spiegel & Watson, 1984), combine with better long-term memory, as found here, to optimize novel word learning in musicians. Such interactions of perceptive and memory functions, as well as transfer effects from musical expertise to long-term memory and novel word learning, add evidence in favor of domain-general networks in the brain underlying perception, cognition, and language processing (for a review on the hotly debated issue of domain-specific vs. domain-general networks in the brain, see Besson, Dittinger, & Barbaroux, 2018). Moreover, because STM and long-term memory are crucial cognitive functions central to language processing, these results highlight the potential impact of music training starting at a young age.

Acknowledgments

We thank all the participants, as well as Chotiga Pattamadilok, Aino Niklas-Salminen, and Mathilde Niklas-Salminen, for registering the auditory stimuli. Moreover, we thank Ken Paller and James McQueen for helpful discussions in the building of the task design. The present work was carried out within The Laboratoire de Neurosciences Cognitives (LNC) Labex Brain and Language Language Institute (ANR-11-LABX-0036) and within the Institute for Language and Communication in the Brain and has been supported by the French government, through the French National Agency for Research (ANR), under the program “Investissements d’Avenir” (ANR-11-IDEX-0001-02). E. D. was supported by a doctoral fellowship from the BLRI (<http://dx.doi.org/10.13039/501100001665>), and B. K. benefited from a grant from the A*MIDEX program (ANR-11-IDEX-0001-02) funded by the French Government under the program “Investissements d’Avenir.”

Reprint requests should be sent to Mireille Besson, Université Publique, Laboratory of Cognitive Neuroscience, Centre Saint-Charles, Pole 3C, Case C, 3 Place Victor Hugo, 13331 Marseille Cedex 3, France, or via e-mail: mireille.besson@univ-amu.fr.

Funding Information

The present work was carried out within the Laboratory of Cognitive Neuroscience (LNC) and was funded by the French government, through the French National Agency for Research (ANR), for the Labex BLRI (ANR-11-LABX-0036) and the program “Investissements d’Avenir” for the Institute for Language and Communication in the Brain (ANR-11-IDEX-0001-02). E. D. was supported by a doctoral fellowship from the BLRI (<http://dx.doi.org/10.13039/501100001665>), and B. K. benefited from a grant from the A*MIDEX program (ANR-11-IDEX-0001-02).

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(an)/M = .408, W(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.

Notes

1. Following phonetic transcription in Thai, 1 refers to low-tone; 0, to mid-tone; p^h, to aspirated voicing; and the colon, to long vowel duration.
2. Pictures were based on the standardized set of 260 pictures built by Snodgrass and Vanderwart (1980) but retraced in Office PowerPoint to ensure sufficient resolution and quality.
3. As for Snodgrass and Vanderwart (1980) pictures, nonobjects from Kroll and Potter (1984) were retraced in Office PowerPoint to ensure sufficient resolution and quality.

4. To simplify, we refer to both Day 1 and Day 2 as “Day 1” (i.e., initial learning session) throughout the paper.
5. “First cycles” refers to the trials of the matching tasks of the first two learning–matching task cycles, and “Final cycles” refers to the trials of the matching tasks of the last two learning–matching task cycles.
6. The number of words to be learned in each experiment was decided based on results of pilot studies.

REFERENCES

- Alario, F. X., & Ferrand, L. (1999). A set of 400 pictures standardized for French: Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. *Behavior Research Methods, Instruments, and Computers*, 31, 531–552. DOI: <https://doi.org/10.3758/bf03200732>, PMID: 10502875
- Asaridou, S. S., & McQueen, J. M. (2013). Speech and music shape the listening brain: Evidence for shared domain-general mechanisms. *Frontiers in Psychology*, 4, 321. DOI: <https://doi.org/10.3389/fpsyg.2013.00321>, PMID: 23761776, PMCID: PMC3671174
- 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. DOI: <https://doi.org/10.1016/j.neuropsychologia.2015.10.020>, PMID: 26476370
- 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. DOI: https://doi.org/10.1162/jocn_a_00013, PMID: 21452941, PMCID: PMC3129368
- Benz, S., Sellaro, R., Hommel, B., & Colzato, L. S. (2016). Music makes the world go round: The impact of musical training on non-musical cognitive functions—A review. *Frontiers in Psychology*, 6, 2023. DOI: <https://doi.org/10.3389/fpsyg.2015.02023>, PMID: 26779111, PMCID: PMC4703819
- Besson, M., Barbaroux, M., & Dittinger, E. (2017). Music in the brain: Music and language processing. In R. Ashley & R. Timmers (Eds.), *The Routledge companion to music cognition*. New York: Routledge.
- Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology*, 2, 94. DOI: <https://doi.org/10.3389/fpsyg.2011.00094>, PMID: 21738519, PMCID: PMC3125524
- Besson, M., Dittinger, E., & Barbaroux, M. (2018). How music training influences language processing: Evidence against informational encapsulation. *Topics in Cognitive Psychology*, 118, 273–288. DOI: <https://doi.org/10.3917/anpsy1.183.0273>
- Besson, M., & Faita, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 1278–1296. DOI: <https://doi.org/10.1037/0096-1523.21.6.1278>
- Bidelman, G. M. (2016). Musicians have enhanced audiovisual multisensory binding: Experience-dependent effects in the double-flash illusion. *Experimental Brain Research*, 234, 3037–3047. DOI: <https://doi.org/10.1007/s00221-016-4705-6>, PMID: 27334887
- Bidelman, G. M., Krishnan, A., & Gandour, J. T. (2011). Enhanced brainstem encoding predicts musicians’ perceptual advantages with pitch. *European Journal of Neuroscience*, 33, 530–538. DOI: <https://doi.org/10.1111/j.1460-9568.2010.07527.x>, PMID: 21198980, PMCID: PMC3059719
- Bidelman, G. M., Weiss, M. W., Moreno, S., & Alain, C. (2014). Coordinated plasticity in brainstem and auditory cortex

- contributes to enhanced categorical speech perception in musicians. *European Journal of Neuroscience*, *40*, 2662–2673. **DOI:** <https://doi.org/10.1111/ejn.12627>, **PMID:** 24890664
- Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer (version 5.2.11). Retrieved from <http://www.praat.org>
- 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. **DOI:** <https://doi.org/10.1080/15475441.2011.614893>, **PMID:** 23125559, **PMCID:** PMC3484686
- Brickenkamp, R., Schmidt-Atzert, L., & Liepmann, D. (2015). *d2-R: Test d'attention concentrée-révisé*. Paris: Éditions Hogrefe France.
- Cardebat, D., Doyon, B., Puel, M., Goulet, P., & Joanette, Y. (1990). Formal and semantic lexical evocation in normal subjects. Performance and dynamics of production as a function of sex, age and educational level. *Acta Neurologica Belgica*, *90*, 207–217. **PMID:** 2124031
- Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–293). Cambridge, MA: MIT Press.
- Chobert, J., Marie, C., François, C., Schön, D., & Besson, M. (2011). Enhanced passive and active processing of syllables in musician children. *Journal of Cognitive Neuroscience*, *23*, 3874–3887. **DOI:** https://doi.org/10.1162/jocn_a_00088, **PMID:** 21736456
- Cunillera, T., Càmarà, E., Toro, J. M., Marco-Pallares, J., Sebastián-Galles, N., Ortiz, H., et al. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *Neuroimage*, *48*, 541–553. **DOI:** <https://doi.org/10.1016/j.neuroimage.2009.06.069>, **PMID:** 19580874
- De Diego Balaguer, R., Toro, J. M., Rodríguez-Fornells, A., & Bachoud-Lévi, A.-C. (2007). Different neurophysiological mechanisms underlying word and rule extraction from speech. *PLoS One*, *2*, e1175. **DOI:** <https://doi.org/10.1371/journal.pone.0001175>, **PMID:** 18000546, **PMCID:** PMC2063512
- Dittinger, E., Barbaroux, M., D'Imperio, M., Jäncke, 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. **DOI:** https://doi.org/10.1162/jocn_a_00997, **PMID:** 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. **DOI:** <https://doi.org/10.3389/fnhum.2017.00233>, **PMID:** 28553213, **PMCID:** PMC5427084
- 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. **DOI:** <https://doi.org/10.1111/ejn.13939>, **PMID:** 29753304
- Dittinger, E., Scherer, J., Jäncke, 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. **DOI:** <https://doi.org/10.1016/j.bandl.2019.104678>, **PMID:** 31450024
- Dobel, C., Junghöfer, M., Breitenstein, C., Klauke, B., Knecht, S., Pantev, C., et al. (2010). New names for known things: On the association of novel word forms with existing semantic information. *Journal of Cognitive Neuroscience*, *22*, 1251–1261. **DOI:** <https://doi.org/10.1162/jocn.2009.21297>, **PMID:** 19583468
- Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science*, *18*, 35–39. **DOI:** <https://doi.org/10.1111/j.1467-9280.2007.01845.x>, **PMID:** 17362375
- Elmer, S., Meyer, M., & Jäncke, L. (2012). Neurofunctional and behavioral correlates of phonetic and temporal categorization in musically trained and untrained subjects. *Cerebral Cortex*, *22*, 650–658. **DOI:** <https://doi.org/10.1093/cercor/bhr142>, **PMID:** 21680844
- Fitzroy, A. B., & Sanders, L. D. (2013). Musical expertise modulates early processing of syntactic violations in language. *Frontiers in Psychology*, *3*, 603. **DOI:** <https://doi.org/10.3389/fpsyg.2012.00603>, **PMID:** 23335905, **PMCID:** PMC3542524
- François, C., Cunillera, T., Garcia, E., Laine, M., & Rodríguez-Fornells, A. (2017). Neurophysiological evidence for the interplay of speech segmentation and word-referent mapping during novel word learning. *Neuropsychologia*, *98*, 56–67. **DOI:** <https://doi.org/10.1016/j.neuropsychologia.2016.10.006>, **PMID:** 27732869
- Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience*, *16*, 1010–1021. **DOI:** <https://doi.org/10.1162/0898929041502706>, **PMID:** 15298788
- George, E. M., & Coch, D. (2011). Music training and working memory: An ERP study. *Neuropsychologia*, *49*, 1083–1094. **DOI:** <https://doi.org/10.1016/j.neuropsychologia.2011.02.001>, **PMID:** 21315092
- Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond “heuristics and biases.” *European Review of Social Psychology*, *2*, 83–115. **DOI:** <https://doi.org/10.1080/14792779143000033>
- Gordon, R. L., Fehd, H. M., & McCandliss, B. D. (2015). Does music training enhance literacy skills? A meta-analysis. *Frontiers in Psychology*, *6*, 1777. **DOI:** <https://doi.org/10.3389/fpsyg.2015.01777>, **PMID:** 26648880, **PMCID:** PMC4664655
- Habibi, A., Damasio, A., Ilari, B., Elliott Sachs, M., & Damasio, H. (2018). Music training and child development: A review of recent findings from a longitudinal study. *Annals of the New York Academy of Sciences*. **DOI:** <https://doi.org/10.1111/nyas.13606>, **PMID:** 29508399
- Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca's region and beyond. *Current Opinion in Neurobiology*, *28*, 136–141. **DOI:** <https://doi.org/10.1016/j.conb.2014.07.013>, **PMID:** 25062474
- Ho, Y.-C., Cheung, M.-C., & Chan, A. S. (2003). Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. *Neuropsychology*, *17*, 439–450. **DOI:** <https://doi.org/10.1037/0894-4105.17.3.439>, **PMID:** 12959510
- Intartaglia, B., White-Schwoch, T., Kraus, N., & Schön, D. (2017). Music training enhances the automatic neural processing of foreign speech sounds. *Scientific Reports*, *7*, 12631. **DOI:** <https://doi.org/10.1038/s41598-017-12575-1>, **PMID:** 28974695, **PMCID:** PMC5626754
- Jaschke, A. C., Honing, H., & Scherder, E. J. A. (2018). Longitudinal analysis of music education on executive functions in primary school children. *Frontiers in Neuroscience*, *12*, 103. **DOI:** <https://doi.org/10.3389/fnins.2018.00103>, **PMID:** 29541017, **PMCID:** PMC5835523
- Jasper, H. H. (1958). The ten–twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology*, *10*, 371–375.
- Kishon-Rabin, L., Amir, O., Vexler, Y., & Zaltz, Y. (2001). Pitch discrimination: Are professional musicians better than non-musicians? *Journal of Basic and Clinical Physiology and Pharmacology*, *12*(2, Suppl), 125–143. **DOI:** <https://doi.org/10.1515/jbcpp.2001.12.2.125>, **PMID:** 11605682
- Koelsch, S., Schröger, E., & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *NeuroReport*, *10*, 1309–1313. **DOI:** <https://doi.org/10.1097/00001756-199904260-00029>, **PMID:** 10363945

- Korkman, M., Kirk, U., & Kemp, S. (2007). *NEPSY-Second Edition (NEPSY-II)*. San Antonio, TX: Harcourt. **DOI:** <https://doi.org/10.1037/t15125-000>
- Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, *11*, 599–605. **DOI:** <https://doi.org/10.1038/nrn2882>, **PMID:** 20648064
- Kraus, N., Strait, D. L., & Parbery-Clark, A. (2012). Cognitive factors shape brain networks for auditory skills: Spotlight on auditory working memory. *Annals of the New York Academy of Sciences*, *1252*, 100–107. **DOI:** <https://doi.org/10.1111/j.1749-6632.2012.06463.x>, **PMID:** 22524346, **PMCID:** PMC3338202
- Kroll, J. F., & Potter, M. C. (1984). Recognizing words, pictures, and concepts: A comparison of lexical, object, and reality decisions. *Journal of Verbal Learning & Verbal Behavior*, *23*, 39–66. **DOI:** [https://doi.org/10.1016/S0022-5371\(84\)90499-7](https://doi.org/10.1016/S0022-5371(84)90499-7)
- 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. **DOI:** <https://doi.org/10.1146/annurev.psych.093008.131123>, **PMID:** 20809790, **PMCID:** PMC4052444
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203–205. **DOI:** <https://doi.org/10.1126/science.7350657>, **PMID:** 7350657
- Kutas, M., Van Petten, C., & Besson, M. (1988). Event-related potential asymmetries during the reading of sentences. *Electroencephalography and Clinical Neurophysiology*, *69*, 218–233. **DOI:** [https://doi.org/10.1016/0013-4694\(88\)90131-9](https://doi.org/10.1016/0013-4694(88)90131-9), **PMID:** 2450003
- Linnavalli, T., Putkinen, V., Lipsanen, J., Huotilainen, M., & Tervaniemi, M. (2018). Music playschool enhances children's linguistic skills. *Scientific Reports*, *8*, 8767. **DOI:** <https://doi.org/10.1038/s41598-018-27126-5>, **PMID:** 29884891, **PMCID:** PMC5993803
- Mankel, K., & Bidelman, G. M. (2018). Inherent auditory skills rather than formal music training shape the neural encoding of speech. *Proceedings of the National Academy of Sciences, U.S.A.*, *115*, 13129–13134. **DOI:** <https://doi.org/10.1073/pnas.1811793115>, **PMID:** 30509989, **PMCID:** PMC6304957
- Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O., & Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *Journal of Cognitive Neuroscience*, *23*, 2701–2715. **DOI:** <https://doi.org/10.1162/jocn.2010.21585>, **PMID:** 20946053
- McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience*, *7*, 703–704. **DOI:** <https://doi.org/10.1038/nn1264>, **PMID:** 15195094
- Mestres-Missé, A., Rodriguez-Fornells, A., & Münte, T. F. (2007). Watching the brain during meaning acquisition. *Cerebral Cortex*, *17*, 1858–1866. **DOI:** <https://doi.org/10.1093/cercor/bhl094>, **PMID:** 17056648
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 227–234. **DOI:** <https://doi.org/10.1037/h0031564>, **PMID:** 5134329
- Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, *219*, 36–47. **DOI:** <https://doi.org/10.1016/j.heares.2006.05.004>, **PMID:** 16839723
- Moreno, S., & Bidelman, G. M. (2014). Examining neural plasticity and cognitive benefit through the unique lens of musical training. *Hearing Research*, *308*, 84–97. **DOI:** <https://doi.org/10.1016/j.heares.2013.09.012>, **PMID:** 24079993
- Moreno, S., & Farzan, F. (2015). Music training and inhibitory control: A multidimensional model. *Annals of the New York Academy of Sciences*, *1337*, 147–152. **DOI:** <https://doi.org/10.1111/nyas.12674>, **PMID:** 25773629
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 519–533. **DOI:** [https://doi.org/10.1016/S0022-5371\(77\)80016-9](https://doi.org/10.1016/S0022-5371(77)80016-9)
- Mosing, M. A., Madison, G., Pedersen, N. L., & Ullén, F. (2016). Investigating cognitive transfer within the framework of music practice: Genetic pleiotropy rather than causality. *Developmental Science*, *19*, 504–512. **DOI:** <https://doi.org/10.1111/desc.12306>, **PMID:** 25939545
- Münte, T. F., Altenmüller, E., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience*, *3*, 473–478. **DOI:** <https://doi.org/10.1038/nrn843>, **PMID:** 12042882
- Paller, K. A., Kutas, M., & Mayes, A. R. (1987). Neural correlates of encoding in an incidental learning paradigm. *Electroencephalography and Clinical Neurophysiology*, *67*, 360–371. **DOI:** [https://doi.org/10.1016/0013-4694\(87\)90124-6](https://doi.org/10.1016/0013-4694(87)90124-6), **PMID:** 2441971
- Paraskevopoulos, E., Kuchenbuch, A., Herholz, S. C., & Pantev, C. (2012). Musical expertise induces audiovisual integration of abstract congruency rules. *Journal of Neuroscience*, *32*, 18196–18203. **DOI:** <https://doi.org/10.1523/JNEUROSCI.1947-12.2012>, **PMID:** 23238733, **PMCID:** PMC6621720
- Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders. The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences*, *999*, 58–75. **DOI:** <https://doi.org/10.1196/annals.1284.006>, **PMID:** 14681118
- Perfetti, C. A., Wlotko, E. W., & Hart, L. A. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1281–1292. **DOI:** <https://doi.org/10.1037/0278-7393.31.6.1281>, **PMID:** 16393047
- Roden, I., Grube, D., Bongard, S., & Kreutz, G. (2014). Does music training enhance working memory performance? Findings from a quasi-experimental longitudinal study. *Psychology of Music*, *42*, 284–298. **DOI:** <https://doi.org/10.1177/0305735612471239>
- Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science*, *15*, 511–514. **DOI:** <https://doi.org/10.1111/j.0956-7976.2004.00711.x>, **PMID:** 15270994
- Schellenberg, E. G. (2011). Examining the association between music lessons and intelligence. *British Journal of Psychology*, *102*, 283–302. **DOI:** <https://doi.org/10.1111/j.2044-8295.2010.02000.x>, **PMID:** 21751987
- Schellenberg, E. G., Corrigall, K. A., Dys, S. P., & Malti, T. (2015). Group music training and children's prosocial skills. *PLoS One*, *10*, e0141449. **DOI:** <https://doi.org/10.1371/journal.pone.0141449>, **PMID:** 26506414, **PMCID:** PMC4624672
- Schulze, K., Dowling, W. J., & Tillmann, B. (2012). Working memory for tonal and atonal sequences during a forward and a backward recognition task. *Music Perception*, *29*, 255–267. **DOI:** <https://doi.org/10.1525/mp.2012.29.3.255>
- Slater, J., Skoe, E., Strait, D. L., O'Connell, S., Thompson, E., & Kraus, N. (2015). Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program. *Behavioural Brain Research*, *291*, 244–252. **DOI:** <https://doi.org/10.1016/j.bbr.2015.05.026>, **PMID:** 26005127
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 174–215.

- DOI:** <https://doi.org/10.1037/0278-7393.6.2.174>, **PMID:** 7373248
- Spiegel, M. F., & Watson, C. S. (1984). Performance on frequency-discrimination tasks by musicians and nonmusicians. *Journal of the Acoustical Society of America*, *76*, 1690–1695. **DOI:** <https://doi.org/10.1121/1.391605>
- Strait, D. L., Kraus, N., Parbery-Clark, A., & Ashley, R. (2010). Musical experience shapes top-down auditory mechanisms: Evidence from masking and auditory attention performance. *Hearing Research*, *261*, 22–29. **DOI:** <https://doi.org/10.1016/j.heares.2009.12.021>, **PMID:** 20018234
- 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. **DOI:** <https://doi.org/10.1016/j.dcn.2015.01.001>, **PMID:** 25660985, **PMCID:** PMC6989776
- Swaminathan, S., Schellenberg, E. G., & Khalil, S. (2017). Revisiting the association between music lessons and intelligence: Training effects or music aptitude? *Intelligence*, *62*, 119–124. **DOI:** <https://doi.org/10.1016/j.intell.2017.03.005>
- Takashima, A., Bakker, I., van Hell, J. G., Janzen, G., & McQueen, J. M. (2014). Richness of information about novel words influences how episodic and semantic memory networks interact during lexicalization. *Neuroimage*, *84*, 265–278. **DOI:** <https://doi.org/10.1016/j.neuroimage.2013.08.023>, **PMID:** 23962957
- Talamini, F., Altoè, G., Carretti, B., & Grassi, M. (2017). Musicians have better memory than nonmusicians: A meta-analysis. *PLoS One*, *12*, e0186773. **DOI:** <https://doi.org/10.1371/journal.pone.0186773>, **PMID:** 29049416, **PMCID:** PMC5648224
- Tervaniemi, M., Castaneda, A., Knoll, M., & Uther, M. (2006). Sound processing in amateur musicians and nonmusicians: Event-related potential and behavioral indices. *NeuroReport*, *17*, 1225–1228. **DOI:** <https://doi.org/10.1097/01.wnr.0000230510.55596.8b>, **PMID:** 16837859
- Torppa, R., Faulkner, A., Laasonen, M., Lipsanen, J., & Sammler, D. (2020). Links of prosodic stress perception and musical activities to language skills of children with cochlear implants and normal hearing. *Ear and Hearing*, *41*, 395–410. **DOI:** <https://doi.org/10.1097/AUD.0000000000000763>, **PMID:** 31397704
- Trainor, L. J., Desjardins, R. N., & Rockel, C. (1999). A comparison of contour and interval processing in musicians and nonmusicians using event-related potentials. *Australian Journal of Psychology*, *51*, 147–153. **DOI:** <https://doi.org/10.1080/00049539908255352>
- Wang, X., Osshner, L., & Reuter-Lorenz, P. A. (2015). Examining the relationship between skilled music training and attention. *Consciousness and Cognition*, *36*, 169–179. **DOI:** <https://doi.org/10.1016/j.concog.2015.06.014>, **PMID:** 26160137
- Wechsler, D. (1997). *Wechsler adult intelligence scale* (3rd ed.). San Antonio, TX: Psychological Corporation. **DOI:** <https://doi.org/10.1037/t49755-000>
- Wojcik, E. H. (2013). Remembering new words: Integrating early memory development into word learning. *Frontiers in Psychology*, *4*, 151. **DOI:** <https://doi.org/10.3389/fpsyg.2013.00151>, **PMID:** 23554599, **PMCID:** PMC3612698
- Wong, P. C. M., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, *28*, 565–585. **DOI:** <https://doi.org/10.1017/S0142716407070312>
- Yang, H., Ma, W., Gong, D., Hu, J., & Yao, D. (2014). A longitudinal study on children's music training experience and academic development. *Scientific Reports*, *4*, 5854. **DOI:** <https://doi.org/10.1038/srep05854>, **PMID:** 25068398, **PMCID:** PMC5376161
- 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. **DOI:** <https://doi.org/10.1371/journal.pone.0099868>, **PMID:** 24937544, **PMCID:** PMC4061064