

ARTICLE

Working memory modulates the effect of music on word learning

Jia Hoong Ong^{1,2*}  and Alice H. D. Chan^{1*}

¹Linguistics & Multilingual Studies, School of Humanities, Nanyang Technological University, Singapore;

²School of Psychology and Clinical Language Sciences, University of Reading, Reading, United Kingdom

*Corresponding authors. Emails: jiahoong.ong@reading.ac.uk; alice@ntu.edu.sg

(Received 26 April 2022; Revised 18 August 2022; Accepted 04 October 2022)

Abstract

While anecdotal evidence suggests that music may facilitate verbal memory, empirical evidence for this is less clear. Here, we examined whether learners' characteristics such as age, working memory (WM), and musical training may influence the effect of music on word learning. Young and older adults learned novel word-referent mappings presented in three music conditions (spoken in the presence of background music, sung in-key, and sung out-of-key) and a control condition (spoken in quiet) and their performance was assessed immediately after learning. We found that whereas age and, to an extent, musical training had a general effect on word learning, WM modulated the effect of music: performance was worse in the music conditions relative to the control condition for learners with lower WM whereas the opposite pattern was observed for those with higher WM. Our results thus highlight the importance of considering individual characteristics in determining the effect of music on verbal memory.

Keywords: word learning; music; song; working memory; individual differences

1. Introduction

Suppose one were to ask an English speaker, “What English letter comes after F?” Chances are the English speaker would start reciting The ABC Song to arrive at the correct response, particularly if that was how they learned the English alphabet. This propensity to retrieve information from memory using songs suggests that information is better remembered and recalled when it is set to a musical melody. It is not surprising, then, that music is often used as a mnemonic device to facilitate learning and remembering information. The present study expands this common intuition by investigating the efficacy of music in language learning; specifically, whether music may be used to support word learning (i.e., the learning of labels for referents).

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike licence (<https://creativecommons.org/licenses/by-nc-sa/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the same Creative Commons licence is included and the original work is properly cited. The written permission of Cambridge University Press must be obtained for commercial re-use.

While anecdotal evidence points to a positive effect of music on learning and memory, empirical evidence for this is less straightforward. Music is typically operationalised in these empirical studies as background music (vs. in quiet) or as a song (vs. speech). Mixed results were obtained from studies comparing memory for items in the presence of background music and in quiet. Some reported a facilitative effect: for instance, learners remembered word lists and second-language vocabulary better when the items were presented with background music compared to those presented in silence (Bottiroli et al., 2014; de Groot, 2006; Kang & Williamson, 2013). Others found that the background music benefit was not seen in item memory (i.e., memory for specific items) but rather in source memory (i.e., memory for the context in which items occur), suggesting that background music provides a facilitative context for encoding information (Ferreri et al., 2014, 2015). On the other hand, some studies reported no difference in recall accuracy between items presented with background music and those presented in silence (de Groot & Smedinga, 2014a; Jäncke et al., 2014; Jäncke & Sandmann, 2010; Nguyen & Grahn, 2017; Reaves et al., 2016). Some even reported worse performance in the background music condition relative to silence, especially in cases where vocal music is used as background music (de Groot & Smedinga, 2014a; Salame & Baddeley, 1989). In an attempt to clarify the inconsistent findings, a meta-analysis revealed that background music appears to have a small detrimental effect on memory in general (Kämpfe et al., 2011). Taken together, the findings on this issue have been highly inconsistent, with the effect of background music on verbal memory ranging from positive to none to negative.

In contrast, studies that compared memory for items that were sung vs. spoken revealed that memory for sung items is typically better than that of spoken (Chazin & Neuschatz, 1990; Ratovohery et al., 2018; Thiessen & Saffran, 2009, but see Racette & Peretz, 2007, for conflicting results). For example, adult learners recalled foreign phrases more accurately when the phrases were sung to them compared to those that were spoken to them (Ludke et al., 2014). Infants, too, appear to remember more from song: 6.5- to 8-month-old infants remembered sung sequences more than spoken sequences as indexed by their ability to differentiate old vs. new sequences (Thiessen & Saffran, 2009). Moreover, the song advantage is reported to be long-lasting, with memory for sung items lasting up to 6 months (Good et al., 2015; Ratovohery et al., 2018). Based on these positive findings, some have explored whether this song advantage may be used with clinical populations such as patients with aphasia (Racette et al., 2006) and patients with mild Alzheimer's disease (Moussard et al., 2012, 2014; Simmons-Stern et al., 2010) and the results have generally been positive with some qualification (e.g., the song advantage tends to be seen after a delay or when sung in unison with an auditory model). While converging evidence suggests a song advantage relative to speech, the advantage is found to be due to certain characteristics of the song itself, such as familiarity of the melody, repetition of melodic structures, and tempo differences between song and speech. Indeed, when these characteristics are controlled, the song advantage tends to disappear (Calvert & Tart, 1993; Kilgour et al., 2000; Wallace, 1994).

Different accounts were put forward to explain the disparate findings in the research on background music and song on verbal memory (Ferreri & Verga, 2016). One such is the 'limited resources' account, which is based on the idea that attentional or cognitive resources are limited (Barrouillet & Camos, 2012; Kahneman, 1973). So, in a situation of a dual task, such as performing a memory task while listening to/processing music, learners' cognitive resources would be

divided between each task, leaving them with less available resources to encode and maintain the target items for the memory task efficiently. Another explanation is the ‘arousal/mood booster’ account (Thompson et al., 2001), in which it is argued that the facilitative effect of music on cognitive tasks is due to an improvement in one’s mood and an increase in one’s arousal level. This would essentially put one in a better state of mind to perform the task. Indeed, some have speculated that auditory stimulation in the background (e.g., white noise or music) may increase dopamine activity, leading to increased attention to perform the task (Angwin et al., 2017). A third account to explain the effect of music on memory is the ‘scaffold/template’ account, which posits that music facilitates one to learn and remember items as it provides a frame to which items can be attached (Purnell-Webb & Speelman, 2008; Thaut et al., 2005). Given that the memory representation of melody and text are integrated to some extent (Serafine et al., 1984), music would thus facilitate by ‘filling in the blanks’ if one should fail to retrieve portions of the verbal material (Ginsborg & Sloboda, 2007). Finally, the findings may also be understood from an informational masking perspective (Eskridge et al., 2012; Scharenborg & Larson, 2018), such that background music, especially ones that have vocals, may mask one’s ability to decode the speech, whereas this would not be the case for sung vocals, in which the speech and music signals are integrated as one.

While this area of research has been studied extensively, there remain several outstanding questions on the influence of the different forms of music on learning and memory, which motivates the present study. Whereas previous studies have compared spoken-in-quiet vs. background music and spoken vs. sung, there has not been any direct comparison between background music vs. sung. Thus, it is unclear whether the two forms of music may have a different effect on word learning. Based on previous findings (i.e., worse performance for background music vs. spoken, and positive or no difference between sung vs. spoken), it is likely that performance on learning sung materials would be superior to learning materials presented with background music. Previous studies have also only used grammatical, in-key music to examine the effect of music on learning and so it is unclear whether music needs to be grammatical to have an effect. With all other things being equal, on the one hand, music that is out-of-key may be a cognitive distraction, costing learners valuable processing resources needed for encoding target items (Slevc et al., 2009). On the other hand, music that is out-of-key note may be considered to be a surprising albeit unpleasant event, especially if it occurs less often than in-key melodies, which may increase its salience and therefore its memorability (Foster & Keane, 2019).

Another outstanding question is what might explain the equivocal findings in the literature. Certainly, the mixed findings are partly attributed to methodological differences across different studies – for example, whether the music used was instrumental or had lyrics, and if so, the language of the lyrics (de Groot & Smedinga, 2014b; Salame & Baddeley, 1989); whether a recognition or a recall task was used (Nguyen & Grahn, 2017); and whether the acoustic differences between the different conditions were properly controlled (Kilgour et al., 2000). Another possibility is related to individual characteristics of the learner, since it appears that not all learners exhibit the same effect of music even within the same study (de Groot, 2006; Küssner et al., 2016). We argue that by not taking these individual characteristics into account, any effect of music on verbal memory may be ‘cancelled out’ when averaged across participants.

Since we conceptualised word learning as a memory task, that is, to commit associations between items (e.g., auditory words and visual forms) to long-term memory (LTM), our general hypothesis is that the influence of music on word

learning (or verbal memory, in general) may be modulated by the learners' cognitive resources and cognitive abilities, assuming all else being equal. That is, those with more resources and enhanced abilities may reap the benefit of music since they have sufficient resources to simultaneously process music and the target items efficiently whereas those with less resources and lower abilities may show the opposite effect (e.g., due to music being distracting, leaving them with less resources to learn target items). In the present study, we operationalised one's cognitive resources as working memory (WM). We propose that WM is a potentially important moderator of the relationship between music and word learning, given its implication in word learning (Baddeley et al., 1998; Martin & Ellis, 2012) and more generally its relation to LTM. For example, according to some WM models, LTM modulates the information held by WM (Jones et al., 2007), whereas according to others, WM is an activated part of the LTM (Cowan, 1999). Thus, it follows that any effect of music that is modulated by WM may invariably also affect word learning.

In addition to WM, we also examined two other individual characteristics that have been shown to affect learners' cognitive resources and/or cognitive abilities that might influence the effect of music on word learning: age and musical training. Older adults are reported to have less cognitive resources and poorer episodic memory than younger adults (Craik & Byrd, 1982; Nilsson, 2003; Rabinowitz et al., 1982) and as such, they tend to show poorer performance in paired-associates task and verbal learning (Korchin & Basowitz, 1957; Meijer et al., 2008; Service & Craik, 1993). In line with our hypothesis that age may moderate the effect of music on word learning, previous studies have demonstrated that older adults tend to be affected by music more so than young adults: whereas there was no effect of music among young adults on verbal memory, older adults tend to show a detrimental effect of background music relative to silence (Reaves et al., 2016) and a facilitative effect of positively valenced familiar song over speech (Ratovohery et al., 2018). Individuals with musical training are reported to have enhanced cognitive abilities including working memory, attention, and executive functioning (Schellenberg & Weiss, 2013; Talamini et al., 2017) than those without musical training. Moreover, they appear to show an aptitude for learning languages, often outperforming their non-musically trained counterparts in various linguistic tasks including word learning (Chobert & Besson, 2013; Dittinger et al., 2016; Kilgour et al., 2000). Thus, given the enhanced cognitive resources/abilities and linguistic abilities associated with musical training, it is possible that musical training may modulate the effect of music on word learning.

In summary, the present study seeks to clarify the mixed findings in the field by examining whether there is an effect of music (operationalised as background music and sung) on word learning relative to spoken-in-quiet. We also explored outstanding questions in the field such as whether there might be differences in word learning performance between items presented in the presence of background music and items that were sung as well as between items sung to melodies that are in-key and those that are out-of-key. Importantly, we test our proposal that the effect of music on word learning may be moderated by individual characteristics such as WM, age, and musical training. Our general hypothesis is that these moderators will influence the direction of the music effect, such that those with higher cognitive resources/abilities will benefit from the effect of music, whereas music will have a detrimental effect on those with lower cognitive resources/abilities.

2. Methods

2.1. Participants

There were two groups of participants: younger and older adults, all of whom were multilingual with English being one of their primary languages. The younger adults were 28 undergraduate students (17 females and 11 males) recruited from a local university. Their age ranged between 20 and 32 ($M = 24.21$, $SD = 3.76$). Some were musically trained¹ ($M = 3.69$, $SD = 4.82$, Range = 0–15 years) and all reported to have normal hearing and normal/corrected-to-normal vision. In addition, participants had their hearing assessed using a pure-tone audiometric screening, and all the younger participants could detect frequency up to 4000 Hz in at least one ear at 25 dB. Participants were screened for their nonverbal intelligence and English receptive vocabulary using Test of Nonverbal Intelligence (TONI, 4th Edition) and Peabody Picture Vocabulary Test (PPVT, 4th Edition), respectively. Their standard scores for both tests were within the normal range (TONI: $M = 104.57$, $SD = 8.37$, Range = 88–124; PPVT: $M = 100.93$, $SD = 11.33$, Range = 81–128).

The older adults were 28 volunteers (18 females and 10 males) recruited from the community whose age ranged between 60 and 87 ($M = 67.11$, $SD = 5.63$). All were reported to have normal hearing as well as normal/corrected-to-normal vision and some were reported to have musical training experience ($M = 1.39$, $SD = 2.95$, Range = 0–10 years). The older adults' hearing was also screened, and all could detect frequency up to 4000 Hz in at least one ear at 40 dB. Participants scored at least 27 on the Mini Mental State Examination (MMSE, 2nd Edition), suggesting that they were cognitively healthy at the time of participating. Their nonverbal intelligence and English receptive vocabulary were within the normal range (TONI: $M = 106.86$, $SD = 10.48$, Range = 90–133; PPVT: $M = 97.79$, $SD = 7.59$, Range = 85–111). Relative to the younger adults, there were no significant differences in their nonverbal intelligence ($t(54) = 0.90$, $p = .371$, $d = 0.24$) and English receptive vocabulary ($t(54) = 1.22$, $p = .228$, $d = 0.33$).

All participants provided their written informed consent prior to participating and they were reimbursed for their participation. The Nanyang Technological University Institutional Review Board approved the study protocol, and all methods were performed in accordance with the relevant guidelines and regulations.

2.2. Stimuli and tasks

2.2.1. Word learning task

The word learning task consisted of 24 word-object pairings. The objects were novel (i.e., they are assumed to be unfamiliar, concrete objects that were not readily named by perceivers) and were taken from The Novel Object and Unusual Name (NOUN)

¹To assess their musical training experience, participants were asked the following questions: "Have you ever had any music lessons (e.g., private or group music lessons) and/or do you practice an instrument on your own?" and "If you answered 'Yes'..., please state the type of lesson you had, the instrument studied, and the year(s) you were enrolled in the lessons, your proficiency on a scale from 1 (not proficient) to 7 (extremely proficient/professional), the number of hours, on average, you played the instrument each week in the past 25 months, and the age at which you started playing the instrument. Please be as specific as possible". From their responses, we counted the duration in years they were enrolled in music lessons as a measure of their musical training experience.

Database (Horst & Hout, 2016). The words were disyllabic pseudowords, which were presented auditorily (see Appendix Table A.1 for the list of pseudowords, and the two languages created for this experiment, i.e., the pseudoword-novel object pairings and the condition to which each pairing was assigned). All auditory stimuli were synthesised using Mac OS X Speech Service with a female voice. Materials for the word learning task can be found at the following link: <https://osf.io/7uywe/>.

The task, presented on PsychoPy (Peirce, 2007), consisted of a learning phase followed by a test phase. In each trial of the learning phase, an object was presented for 4.5 s in the middle of screen. 1.5 s after the object appears, a pseudoword for the object was presented auditorily in a carrier phrase “This is a [pseudoword]”. Importantly, the carrier phrase and pseudowords were presented in four within-subject conditions – spoken-in-quiet condition (Spoken) and three music conditions (spoken in the presence of background music (Bg), sung in-key (Sung-in), and sung out-of-key (Sung-out)), all to the tune of novel five-note melodies – with each condition consisting of six word-object pairings. In the Spoken condition, the carrier phrase and pseudoword were spoken with a natural prosody in quiet. In the Bg condition, each spoken sentence was accompanied by a soft sine-tone melody (with its amplitude peak scaled to 0.2, or approximately 77 dB SPL, on Praat; Boersma & Weenink, 2013) in which each tone was linked to a syllable and their duration matched. In the Sung conditions, the pitch of each spoken syllable of the carrier phrase was manipulated on Praat (Boersma & Weenink, 2013) such that it would be ‘sung’ to tones implying a major key (e.g., “This is a...” sung to E-G-C, which implies the key of C). Specifically, we levelled the pitch contour of each syllable and shifted its pitch to a particular frequency (tuned to A4 = 440 Hz). The pitch of the pseudowords was also manipulated in the same way, resulting in sung syllables that would either be in-key (Sung-in) or out-of-key (Sung-out) relative to implied key of the carrier phrase. For example, relative to a carrier in the key of C (E-G-C), an in-key pseudoword ‘rin-ba’ would be sung to D-C (which is consistent with the key of C) whereas an out-of-key pseudoword ‘lu-gash’ would be sung to G#-F# (which implies the key of F#). We defined ‘out-of-key’ as being the key directly opposite the implied key of the carrier phrase on the circle of fifths, a representation of the relationship between musical keys in music theory. This manipulation should theoretically be the most out-of-key possible, and thus the key violation to the implied musical key should be obvious to the adult listeners given that key membership is learned by the age of 5 years old (Trainor & Hannon, 2013). To avoid potential confounding effects of pseudoword-object pairings and their allocation to a condition, we created two languages, which contained the same stimuli but we randomly paired the pseudowords and the objects and their allocation to one of the four conditions in both languages (see Appendix Table A.1). Participants were randomly assigned to one of the two languages at the start of the experiment. We focused on pitch, rather than rhythm, manipulations to differentiate the conditions given the relatively short utterance on every trial (5 syllables/notes). The duration of the stimuli was similar between the conditions: all had the same duration of the carrier phrase (786 ms) plus the duration of the disyllabic word, which ranged between 349 and 669 ms ($M = 508$ ms, $SD = 68$ ms). A 2 (List) \times 4 (Condition) ANOVA on the duration of the pseudowords revealed no significant main effects of List ($F(1, 40) = 0.46$, $p = .503$) and Condition ($F(3, 40) = 0.32$, $p = .812$), nor a significant interaction between the two ($F(3, 40) = .26$, $p = .854$). Each word-object pairing was presented twice during the learning phase (with an inter-trial interval of 500 ms) in a

pseudorandomised order such that the conditions were not blocked and that no two consecutive learning trials were the same.

Directly after the learning phase, participants completed a six-alternative forced-choice recognition test. On every test trial, participants heard a pseudoword (all of which were spoken) and they had to choose which of the six images on the screen corresponded to the pseudoword. The target image was presented with five distractor images, which were taken from the same condition (e.g., if the target image was from Sung-in condition, then all the distractor images were also from Sung-in condition). Each pairing was tested once in a randomised order.

2.2.2. Auditory working memory task

To measure participants' working memory, we used the Auditory Working Memory subtest from Woodcock–Johnson Tests of Cognitive Abilities (3rd Edition). Participants heard a list of objects and digits in a random order on each trial. At the end of every list, participants had to first recall the objects in the order that the objects were presented and then the digits in the order that the digits were presented. The number of words in each list ranged between three and eight. Two points were given to each list when participants correctly recalled the objects and the digits in the correct order. One point was given when they correctly recalled either the objects or the digits in order. The maximum raw score for this task is 42.

2.3. Procedure

Participants completed all the tasks in a single session in the following fixed order: word learning, TONI, PPVT, auditory working memory, and demographic questionnaire. Additionally, the older adults completed MMSE at the start of the experiment. Prior to the word learning task, participants adjusted their headphone volume to a comfortable level to ensure they could hear the utterances clearly.

2.4. Data analysis

Independent samples *t*-tests were first conducted to examine whether younger and older adults differed in their working memory scores and musical training experience.

Data of the word learning task were fitted using mixed effects logistic regression using the *lme4* package version 1.1.27.1 (Douglas et al., 2015) in R version 4.1.2 (R Core Team, 2021). Data and the analysis script for the experiments reported in this manuscript are available at the following link: <https://osf.io/7uywe/>. The analysis choice is motivated by the binary nature of the dependent variable, Accuracy (Jaeger, 2008). There were four predictors: Condition (Helmert coded to test three planned comparisons: (i) Spoken vs. the three music conditions collapsed; (ii) Bg vs. the two sung conditions collapsed; and (iii) Sung In-Key vs. Sung Out-of-Key); Age Group (Age; effect-coded: Younger vs Older); Working Memory (WM; centred continuous variable by Age Group); and Years of Musical Training (Training; centred continuous variable by Age Group). We tested our a priori contrasts for Condition rather than an omnibus test of Condition given that the former approach is more powerful and appropriate (Schad et al., 2020). We modelled the data with all the predictors and the two-way interactions between Condition and Age, Condition and WM, and

Condition and Training, to examine whether these factors modulate the effect of music on word learning. Pairwise comparisons, if any, were conducted using *emmeans* package version 1.7.1.1. As random effects, we entered random by-subject and by-item intercepts and random by-subject slopes for Condition.

3. Results

Independent samples *t*-tests revealed that the older adults had lower auditory WM raw scores ($t(54) = 8.26, p < .001, d = 2.21$) and less years of musical training ($t(54) = 2.15, p = .036, d = 0.57$) than the younger adults. In the subsequent analysis, we included WM scores and years of musical training, both of which were centred by age group.

Descriptive statistics of the overall proportion of correct responses for the word learning task by age group and condition are displayed in Table 1. The overall performance for both age groups was relatively high, given that the 95% confidence intervals were well above the chance level for a 6-alternative forced-choice task (i.e., 1/6, or 0.167).

The model summary of the word learning data can be found in Appendix Table A.2. Findings from the model suggest that younger adults generally performed better than older adults ($\beta = -0.27, SE = 0.14, z = 1.97, p = .048$), and that there was a marginal positive relationship between word learning performance and years of musical training ($\beta = 0.07, SE = 0.04, z = 1.93, p = .053$). There was no influence of music on word learning generally, that is, the performance on each of the conditions was not significantly different from each other (see Fig. 1).

The model also examined whether the influence of music on word learning was modulated by individual characteristics (i.e., the two-way interactions). We found that while age and years of musical training did not modulate the effect of music (i.e., none of the interactions involving age and musical training were significant), working memory (WM) modulated the effect of Spoken condition relative to the music conditions (i.e., Spoken vs. Music \times WM interaction: $\beta = -0.02, SE = 0.01, z = -2.13, p = .033$; see Fig. 2). The estimated effect (i.e., the slope) of WM on spoken was significantly more negative than the music conditions ($B_{diff} = -0.266, SE = 0.12, z = -2.13, p = .033$). Descriptively, those with lower WM performed worse on the music conditions relative to the Spoken condition whereas those with higher WM scored higher on the music conditions compared to the Spoken condition. No other interactions involving WM were significant.

Table 1. Descriptive statistics (mean, 95% confidence intervals (CI), and standard deviation) of the overall proportion correct for the word learning task by age group and condition (M, SD, CI)

Condition	Younger adults ($n = 28$)		Older adults ($n = 28$)	
	Mean (95% CI)	SD	Mean (95% CI)	SD
Spoken	0.65 [0.55, 0.76]	0.29	0.52 [0.43, 0.62]	0.25
Bg	0.63 [0.51, 0.74]	0.31	0.55 [0.45, 0.65]	0.27
Sung-In	0.72 [0.63, 0.81]	0.24	0.57 [0.45, 0.68]	0.3
Sung-Out	0.68 [0.58, 0.78]	0.27	0.61 [0.51, 0.72]	0.28

Note. Spoken = Spoken-in-quiet condition; Sung-in = Sung in-key condition; Sung out = Sung-out-of-key condition; Bg = Spoken in the presence of background music condition.

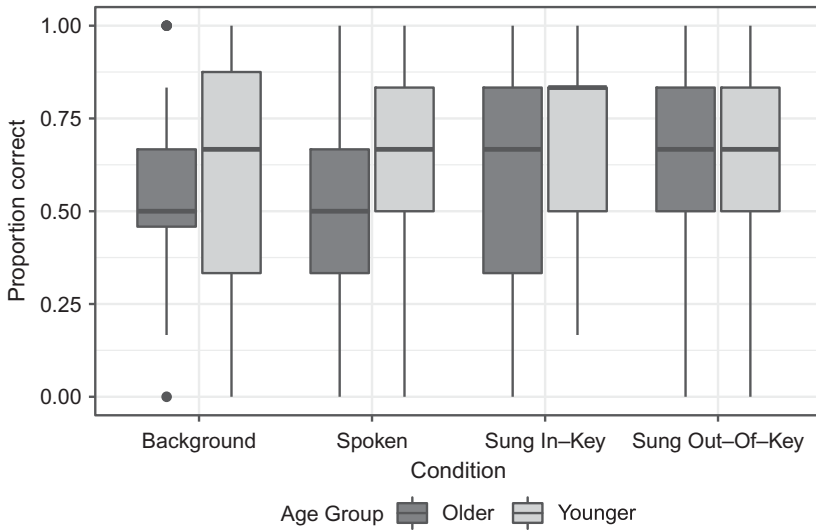


Fig. 1. Performance on each of the conditions in the word learning task by age group.

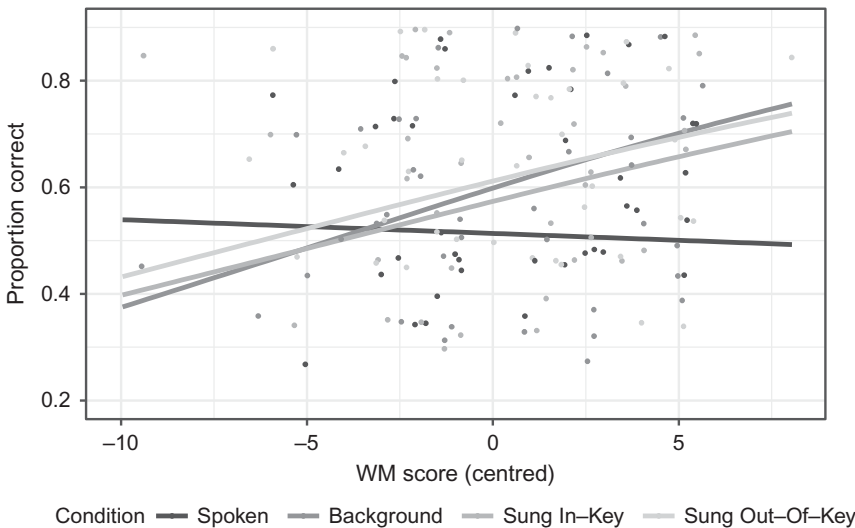


Fig. 2. Predicted word learning performance on the Spoken condition (in dark grey) vs. the three music conditions (Background, Sung In-Key, and Sung Out-of-Key, in increasing lighter shades of grey) by centred Working Memory (WM) scores using fitted data. Individual data points are also plotted as dots.

4. Discussion

The present study clarifies what effect music has on memory, from a word-learning perspective, given mixed findings in previous studies. We compared word learning performance in three music conditions – background music, sung in-key, and sung out-of-key – with a control condition, that is, spoken in quiet. We also explored

several outstanding questions in the field, such as whether there would be any difference in verbal memory for items presented with background music vs. sung items, and whether items sung in-key would be remembered differently than that sung out-of-key. Crucially, we investigated these questions from an individual differences perspective, by examining whether individual characteristics such as age, working memory (WM), and musical training may modulate the effects of music on word learning. We argue that these individual characteristics may partly contribute to the mixed findings in the literature, given that any effect of music (either positive or negative) may be 'cancelled out' by not taking them into consideration.

We found that in general, when individual characteristics were not taken into consideration, there was no effect of music on word learning relative to spoken-in-quiet and no difference in performance between background music and sung conditions and between sung in-key and sung out-of-key conditions. Extending previous studies that have typically compared either speech and background music or speech and song, our study demonstrated that the various music conditions do not differentially affect verbal memory. This is somewhat surprising, given that previous studies have found similar or worse performance for background music vs. spoken (e.g., Jäncke & Sandmann, 2010; Kämpfe et al., 2011; Reaves et al., 2016) while superior performance for sung vs. spoken has been reported (e.g., Ludke et al., 2014; Ratovohery et al., 2018). So, one might expect that performance would be better in the sung conditions than in the background music condition. We suspect that the lack of a difference between the music conditions may be due to our manipulation methods. We wanted the music conditions to be as similar as possible to one another and to the spoken condition. Thus, the typical advantages afforded by songs such as repetition of melodic structures and slower articulation or tempo were not present in our sung conditions. Moreover, our use of short five-note melodies may not be 'surprising' or memorable enough when it is out-of-key for the learners to show differential performances between the two sung conditions.

Our results indicated that across conditions young adults performed better than older adults, unsurprisingly, given that older adults tend to have poorer declarative memory (Korchin & Basowitz, 1957; Meijer et al., 2008; Service & Craik, 1993). There was also a marginal positive relationship between word learning performance and years of musical training across conditions, which is also expected given that musical training appears to enhance general cognitive abilities and is related to higher linguistic aptitude (Chobert & Besson, 2013; Dittinger et al., 2016; Kilgour et al., 2000; Schellenberg & Weiss, 2013; Talamini et al., 2017). The older adults generally had less musical training experience than the younger adults in our sample, and so while we took that into consideration in our model (i.e., we mean-centred their musical experience by age group), it is possible that the effect of musical training may be stronger if there were more musically trained older adults in our sample.

Beyond the general effects of age and musical training on word learning, there was no evidence that the effect of music was modulated by those individual characteristics. Thus, our results appear to not be in line with those that found an effect of music among older adults but not younger adults (Ratovohery et al., 2018; Reaves et al., 2016). The difference in findings is likely due to the type of music used – whereas previous studies typically employed rich, emotionally charged music, our study has opted for a more acoustically controlled, shorter melodies in order to be comparable to the control/spoken-in-quiet condition. In contrast to age and musical training, we found evidence of WM modulating the effect of music on word learning.

Specifically, words that were spoken in the presence of background music and words that were sung were learned better than words spoken in quiet by those with higher WM whereas the opposite pattern was observed among those with lower WM. Based on the cognitive account discussed previously, we propose that music listening may involuntarily draw one's cognitive resources, leaving listeners with lower WM with insufficient resources to properly encode the word-referent mappings. On the other hand, those with higher WM would still have sufficient resources for efficient encoding of the word-referent mappings, which leads to their ability to reap the benefits of music in learning.

Several limitations of the study are worth noting. Our sample size, though comparable to other studies (Jäncke & Sandmann, 2010; Ratovohery et al., 2018), is modest. With a larger sample size, it would be possible to examine whether different forms of musical training (e.g., vocalists vs. instrumental musicians) may affect performances in the various music conditions. Our measure of musical experience is relatively crude and could benefit from using more sophisticated measures, such as a point-system differentiating type of musical training (e.g., private vs. group; Russo et al., 2007) or standardised questionnaires such as the Goldsmiths Musical Sophistication Index (Gold-MSI; Müllensiefen et al., 2014). Finally, we did not screen our participants for amusia. Though rare (occurring in about 1.5% of the general population; Peretz & Vuvar, 2017) and the selected out-of-key melodies were the most out-of-key according to the circle of fifths (i.e., we selected the key opposite the implied key), it is possible that the amusics, if there were any, might not have been able to detect the implied keys, which may affect the results. This should be considered in future studies.

This study provides evidence that a learner's WM influences whether music exerts an effect on word learning. Given these findings, it appears that the effect of music on verbal memory is partly cognitive in nature. Moreover, we suggest that the equivocal results in previous studies may be partly due to not taking this and possibly other characteristics into consideration, which would average out any potential effects of music at the group level. It should be noted that previous studies have used different methodologies (e.g., the modality of the target items (e.g., printed words); background music not following the presentation rate of spoken material, etc.), which may involve different mechanisms and memory components than our current paradigm, and so this may partly explain the contrastive findings with the present study. Nonetheless, our present study has shown that individual characteristics do seem to influence the effect of music and taking this into account will provide us with a more nuanced understanding of how music may influence our memory. Further work is needed to identify other characteristics that may similarly exert an influence, which has clear implications for education (e.g., whether the effect may be stronger among learners with tone language experience given their extensive pitch experience), and how these characteristics may interact with various aspects of the music (e.g., familiarity with and likability of melodies) and methodology (e.g., having the same mode of presentation during learning and test). From a theoretical perspective, it would be interesting to examine whether music may exert an influence beyond word learning, such as grammar learning. According to certain language models (Hamrick et al., 2018; Ullman, 2004), the same memory component (i.e., declarative memory) is implicated in vocabulary learning and learning grammar of a second language in the initial stages. Thus, if music affects the retrieval of information from

declarative memory, then based on our present findings, music may similarly influence initial grammar learning.

5. Conclusions

In conclusion, we found that WM appears to modulate the effect of music on word learning such that those with higher WM appear to benefit from the effect of music whereas those with lower WM were disadvantaged by the presence of music. More generally, our results highlight the need to consider individual characteristics in determining the effect of music. Indeed, we propose that the mixed results seen in previous studies may be partly due to the effect of music being ‘averaged out’ from not taking into account learner’s characteristics. For pedagogical and educational implications, further studies are necessary to identify other characteristics that may similarly influence the effect of music, and to explore whether the effect of music extends beyond word learning to other aspects of language acquisition.

Acknowledgements. We would like to thank all the participants who volunteered their time and Cheryl Choo and Hannah Lim for assistance with data collection.

Funding. This work was supported by the Ministry of Education (Singapore) Social Science Research Thematic Grant (SSRTG) (grant number: MOE2019-SSRTG-016), Academic Research Fund (MoE AcRF) Tier 2 (grant number MOE2019-T2-1-125), and Tier 1 (grant number RG71/18) awarded to AHDC. JHO was supported by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 887283.

Data availability statement. The dataset supporting the conclusions of this article is available at <https://osf.io/7uywe/>.

References

- Angwin, A. J., Wilson, W. J., Arnott, W. L., Signorini, A., Robert, J., & Copland, D. A. (2017). White noise enhances new-word learning in healthy adults. *Scientific Reports*, 7, 1–6. <https://doi.org/10.1038/s41598-017-13383-3>
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105(1), 158–173. <https://doi.org/10.1037/0033-295X.105.1.158>
- Barrouillet, P., & Camos, V. (2012). The time-based resource-sharing model of working memory. In N. Osaka, R. H. Logie, & M. D’Esposito (Eds.), *The cognitive neuroscience of working memory* (pp. 59–80). Oxford Scholarship Online. <https://doi.org/10.1093/acprof:oso/9780198570394.003.0004>
- Boersma, P., & Weenink, D. (2013). *Praat: Doing phonetics by computer* (5.3.42).
- Bottiroli, S., Rosi, A., Russo, R., Vecchi, T., & Cavallini, E. (2014). The cognitive effects of listening to background music on older adults: Processing speed improves with upbeat music, while memory seems to benefit from both upbeat and downbeat music. *Frontiers in Aging Neuroscience*, 6(284), 1–7. <https://doi.org/10.3389/fnagi.2014.00284>
- Calvert, S. L., & Tart, M. (1993). Song versus verbal forms for very-long-term, long-term, and short-term verbatim recall. *Journal of Applied Developmental Psychology*, 14(2), 245–260. [https://doi.org/10.1016/0193-3973\(93\)90035-T](https://doi.org/10.1016/0193-3973(93)90035-T)
- Chazin, S., & Neuschatz, J. S. (1990). Using a mnemonic to aid in the recall of unfamiliar information. *Perceptual and Motor Skills*, 1067–1071. <https://doi.org/10.2466/pms.1990.71.3f.1067>
- Chobert, J., & Besson, M. (2013). Musical expertise and second language learning. *Brain Sciences*, 3(2), 923–940. <https://doi.org/10.3390/brainsci3020923>
- Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge University Press. <https://doi.org/10.1017/CBO9781139174909.006>

- Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik & S. E. Trehub (Eds.), *Aging and Cognitive Processes*. Plenum Publishing Corporation. <https://doi.org/10.1007/978-1-4684-4178-9>.
- de Groot, A. M. B. (2006). Effects of stimulus characteristics and background music on foreign language vocabulary learning and forgetting. *Language Learning*, 56(3), 463–506. <https://doi.org/10.1111/j.1467-9922.2006.00374.x>
- de Groot, A. M. B., & Smedinga, H. E. (2014a). Let the music play! A short term but no long-term detrimental effect of vocal background music with familiar language lyrics on foreign language vocabulary learning. *Studies in Second Language Acquisition*, 36, 681–707. <https://doi.org/10.1017/S0272263114000059>
- 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(10), 1584–1602. https://doi.org/10.1162/jocn_a_00997
- Douglas, B., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Eskridge, E. N., Galvin, J. J., Aronoff, J. M., Li, T., & Fu, Q.-J. (2012). Speech perception with music maskers by cochlear implant users and normal hearing listeners. *Journal of Speech, Language, and Hearing Research : JSLHR*, 55(3), 800–810. [https://doi.org/10.1044/1092-4388\(2011/11-0124](https://doi.org/10.1044/1092-4388(2011/11-0124)
- Ferreri, L., Bigand, E., & Bugaiska, A. (2015). The positive effect of music on source memory. *Musicae Scientiae*, 19(4), 402–411. <https://doi.org/10.1177/1029864915604684>
- Ferreri, L., Bigand, E., Perrey, S., Muthalib, M., Bard, P., & Bugaiska, A. (2014). Less effort, better results: How does music act on prefrontal cortex in older adults during verbal encoding? An fNIRS study. *Frontiers in Human Neuroscience*, 8(May), 1–11. <https://doi.org/10.3389/fnhum.2014.00301>
- Ferreri, L., & Verga, L. (2016). Benefits of music on verbal learning and memory: How and when does it work? *Music Perception*, 34(2), 167–182. <https://doi.org/10.1525/mp.2016.34.2.167>
- Foster, M. I., & Keane, M. T. (2019). The role of surprise in learning: Different surprising outcomes affect memorability differentially. *Topics in Cognitive Science*, 11(1), 75–87. <https://doi.org/10.1111/tops.12392>
- Ginsborg, J., & Sloboda, J. A. (2007). Singers' recall for the words and melody of a new, unaccompanied song. *Psychology of Music*, 35(3), 421–440. <https://doi.org/10.1177/0305735607072654>
- Good, A. J., Russo, F. A., & Sullivan, J. (2015). The efficacy of singing in foreign-language learning. *Psychology of Music*, 43(5), 627–640. <https://doi.org/10.1177/0305735614528833>
- Hamrick, P., Lum, J. A. G., & Ullman, M. T. (2018). Child first language and adult second language are both tied to general-purpose learning systems. *Proceedings of the National Academy of Sciences*, 1–6. <https://doi.org/10.1073/pnas.1713975115>
- Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behavior Research Methods*, 48, 1393–1409. <https://doi.org/10.3758/s13428-015-0647-3>
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59(4), 434–446. <https://doi.org/10.1016/j.jml.2007.11.007>.
Categorical
- Jäncke, L., Brügger, E., Brummer, M., Scherrer, S., & Alahmadi, N. (2014). Verbal learning in the context of background music: No influence of vocals and instrumentals on verbal learning. *Behavioral and Brain Functions*, 10, 10. <https://doi.org/10.1186/1744-9081-10-10>
- Jäncke, L., & Sandmann, P. (2010). Music listening while you learn: No influence of background music on verbal learning. *Behavioral and Brain Functions : BBF*, 6, 3. <https://doi.org/10.1186/1744-9081-6-3>
- Jones, G., Gobet, F., & Pine, J. M. (2007). Linking working memory and long-term memory: A computational model of the learning of new words. *Developmental Science*, 10(6), 853–873. <https://doi.org/10.1111/j.1467-7687.2007.00638.x>
- Kahneman, D. (1973). *Attention and Effort*. Prentice-Hall.
- Kämpfe, J., Sedlmeier, P., & Renkewitz, F. (2011). The impact of background music on adult listeners: A meta-analysis. *Psychology of Music*, 39(4), 424–448. <https://doi.org/10.1177/0305735610376261>
- Kang, H. J., & Williamson, V. J. (2013). Background music can aid second language learning. *Psychology of Music*, 42(5), 728–747. <https://doi.org/10.1177/0305735613485152>
- Kilgour, A. R., Jakobson, L. S., & Cuddy, L. L. (2000). Music training and rate of presentation as mediators of text and song recall. *Memory & Cognition*, 28(5), 700–710. <https://doi.org/10.3758/BF03198404>

- Korchin, S. J., & Basowitz, H. (1957). Age differences in verbal learning. *The Journal of Abnormal and Social Psychology*, 54(1), 64–69. <https://doi.org/10.1037/h0049240>
- Küssner, M. B., De Groot, A. M. B., Hofman, W. F., & Hillen, M. A. (2016). EEG beta power but not background music predicts the recall scores in a foreign-vocabulary learning task. *PLoS ONE*, 11(8), 1–17. <https://doi.org/10.1371/journal.pone.0161387>
- Ludke, K. M., Ferreira, F., & Overy, K. (2014). Singing can facilitate foreign language learning. *Memory & Cognition*, 42(July 2013), 41–52. <https://doi.org/10.3758/s13421-013-0342-5>
- Martin, K. I., & Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition*, 7(62), 379–413. <https://doi.org/10.1017/S0272263112000125>
- Meijer, W. A., de Groot, R. H. M., Van Boxtel, M. P. J., Van Gerven, P. W. M., & Jolles, J. (2008). Are age differences in verbal learning related to interstimulus interval and education? *Experimental Aging Research*, 34(4), 323–339. <https://doi.org/10.1080/03610730802273910>
- Moussard, A., Bigand, E., Belleville, S., & Peretz, I. (2012). Music as an aid to learn new verbal information in Alzheimer's Disease. *Music Perception*, 29(5), 521–532. <https://doi.org/10.1525/MP.2012.29.5.521>
- Moussard, A., Bigand, E., Belleville, S., & Peretz, I. (2014). Learning sung lyrics aids retention in normal ageing and Alzheimer's disease. *Neuropsychological Rehabilitation*, 24(6), 894–917. <https://doi.org/10.1080/09602011.2014.917982>
- Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PLoS ONE*, 9(2), e89642. <https://doi.org/10.1371/journal.pone.0089642>
- Nguyen, T., & Grahn, J. A. (2017). Mind your music: The effects of music-induced mood and arousal across different memory tasks. *Psychomusicology: Music, Mind, and Brain*, 27(2), 81–94. <https://doi.org/10.1037/pmu0000178>
- Nilsson, L.-G. (2003). Memory function in normal aging. *Acta Neurologica Scandinavica*, 107(s179), 7–13. <https://doi.org/10.1034/j.1600-0404.107.s179.5.x>
- Peirce, J. W. (2007). PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1–2), 8–13. <https://doi.org/10.1016/j.jneumeth.2006.11.017>
- Peretz, I., & Vuvan, D. T. (2017). Prevalence of congenital amusia. *European Journal of Human Genetics*, 25(5), 625–630. <https://doi.org/10.1038/ejhg.2017.15>
- Purnell-Webb, P., & Speelman, C. P. (2008). Effect of music on memory for text. *Perceptual and Motor Skills*, 106(3), 927–957.
- R Core Team. (2021). *R: A language and environment for statistical computing* (4.1.2). R Foundation for Statistical Computing.
- Rabinowitz, J. C., Craik, F. I. M., & Ackerman, B. P. (1982). A processing resource account of age differences in recall. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, 36(2), 325–344. <https://doi.org/10.1037/h0080643>
- Racette, A., Bard, C., & Peretz, I. (2006). Making non-fluent aphasics speak: Sing along! *Brain*, 129(10), 2571–2584. <https://doi.org/10.1093/brain/awl250>
- Racette, A., & Peretz, I. (2007). Learning lyrics: To sing or not to sing? *Memory & Cognition*, 35(2), 242–253. <https://doi.org/10.3758/BF03193445>
- Ratovohery, S., Baudouin, A., Gachet, A., Palisson, J., & Narme, P. (2018). Is music a memory booster in normal aging? The influence of emotion. *Memory*, 1–11. <https://doi.org/10.1080/09658211.2018.1475571>
- Reaves, S., Graham, B., Grahn, J., Rabannifard, P., & Duarte, A. (2016). Turn off the music! Music impairs visual associative memory performance in older adults. *The Gerontologist*, 56(3), 569–577. <https://doi.org/10.1093/geront/gnu113>
- Russo, F. A., Cuddy, L. L., Galembo, A., & Thompson, W. F. (2007). Sensitivity to tonality across the pitch range. *Perception*, 36(5), 781–790. <https://doi.org/10.1068/p5435>
- Salame, P., & Baddeley, A. (1989). Effects of background music on phonological short-term memory. *The Quarterly Journal of Experimental Psychology*, 41(1), 107–122. <https://doi.org/10.1080/14640748908402355>
- Schad, D. J., Vasisht, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, 110, 104038. <https://doi.org/10.1016/j.jml.2019.104038>

- Scharenborg, O., & Larson, M. (2018). The conversation continues: The effect of lyrics and music complexity of background music on spoken-word recognition. *Interspeech*, 2018, 2280–2284. <https://doi.org/10.21437/Interspeech.2018-1088>
- Schellenberg, E. G., & Weiss, M. W. (2013). Music and cognitive abilities. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 499–550). Elsevier. <https://doi.org/10.1016/B978-0-12-381460-9.00012-2>
- Serafine, M. L., Crowder, R. G., & Repp, B. H. (1984). Integration of melody and text in memory for songs. *Cognition*, 16(3), 285–303. [https://doi.org/10.1016/0010-0277\(84\)90031-3](https://doi.org/10.1016/0010-0277(84)90031-3)
- Service, E., & Craik, F. I. M. (1993). Differences between young and older adults in learning a foreign vocabulary. *Journal of Memory and Language*, 32, 608–623.
- Simmons-Stern, N. R., Budson, A. E., & Ally, B. A. (2010). Music as a memory enhance in patients with Alzheimer's Disease. *Neuropsychologia*, 48(10), 3164–3167. <https://doi.org/10.1016/j.neuropsychologia.2010.04.033>. Music
- Slevc, L. R., Rosenberger, J. C., & Patel, A. D. (2009). Making psycholinguistics musical: Self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychonomic Bulletin & Review*, 16(2), 374–381.
- Talamini, F., Altoè, G., Carretti, B., & Grassi, M. (2017). Musicians have better memory than nonmusicians: A meta-analysis. *PLoS ONE*, 12(10), 1–21. <https://doi.org/10.1371/journal.pone.0186773>
- Thaut, M. H., Peterson, D. A., & McIntosh, G. C. (2005). Temporal entrainment of cognitive functions: Musical mnemonics induce brain plasticity and oscillatory synchrony in neural networks underlying memory. *Annals of the New York Academy of Sciences*, 1060, 243–254. <https://doi.org/10.1196/annals.1360.017>
- Thiessen, E. D., & Saffran, J. R. (2009). How the melody facilitates the message and vice versa in infant learning and memory. *Annals of the New York Academy of Sciences*, 1169, 225–233. <https://doi.org/10.1111/j.1749-6632.2009.04547.x>
- Thompson, W. F., Schellenberg, E. G., & Husain, G. (2001). Arousal, mood, and the Mozart effect. *Psychological Science*, 12(3), 248–251.
- Trainor, L. J., & Hannon, E. E. (2013). Musical development. *The Psychology of Music* (Issue 1997). <https://doi.org/10.1016/B978-0-12-381460-9.00011-0>
- Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, 92(1–2), 231–270. <https://doi.org/10.1016/j.cognition.2003.10.008>
- Wallace, W. T. (1994). Memory for music: Effect of melody on recall of text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(6), 1471–1485. <https://doi.org/10.1037/0278-7393.20.6.1471>

Appendices

Table A.1. The pairings of the pseudoword and the object in the two languages to which participants were randomly assigned at the start of the experiment. Note that the pseudowords are identical in both languages but may differ in their pairings with the object and the condition allocation. Objects are arbitrarily identified using their code in this table

Language A			Language B		
Pseudoword	Object	Condition	Pseudoword	Object	Condition
Halloon	P057	Spoken	Fougan	P043	Spoken
Rinba	P064	Spoken	Lencal	P046	Spoken
Triment	P054	Spoken	Derson	P041	Spoken
Derson	P044	Spoken	Mirton	P052	Spoken
Bevlam	P045	Spoken	Honpo	P057	Spoken
Lencal	P062	Spoken	Chadew	P045	Spoken
Lugash	P061	Sung-in	Maver	P051	Sung-in
Chottle	P048	Sung-in	Groonda	P055	Sung-in
Groonda	P046	Sung-in	Frinkle	P050	Sung-in
Fougan	P059	Sung-in	Triment	P047	Sung-in
Frinkle	P042	Sung-in	Warris	P063	Sung-in
Maver	P051	Sung-in	Dutran	P042	Sung-in
Wuggle	P050	Sung-out	Gruton	P054	Sung-out
Honpo	P041	Sung-out	Halloon	P060	Sung-out
Plinka	P056	Sung-out	Rinba	P062	Sung-out
Dutran	P049	Sung-out	Bodwit	P064	Sung-out
Pooda	P053	Sung-out	Lugash	P058	Sung-out
Gruton	P055	Sung-out	Pooda	P053	Sung-out
Tizuk	P058	Bg	Plinka	P049	Bg
Bodwit	P047	Bg	Rihaz	P059	Bg
Mirton	P052	Bg	Bevlam	P056	Bg
Rihaz	P043	Bg	Wuggle	P048	Bg
Chadew	P063	Bg	Tizuk	P061	Bg
Warris	P060	Bg	Chottle	P044	Bg

Note. Spoken = Spoken-in-quiet condition; Sung-in = Sung in-key condition; Sung out = Sung-out-of-key condition; Bg = Spoken in the presence of background music condition.

Table A.2. Output of the mixed effects logistic regression model

Predictor	β [95% CI]	SE	z	p
Intercept	0.57 [0.26, 0.88]	0.16	3.60	<.001
Spoken vs. Music	-0.06 [-0.16, 0.03]	0.05	-1.39	.166
Bg vs. Sung	-0.02 [-0.16, 0.11]	0.07	-0.35	.730
Sung-in vs. Sung-out	0.03 [-0.25, 0.31]	0.14	0.20	.842
WM	0.06 [-0.02, 0.13]	0.04	1.54	.123
Age	-0.27 [-0.53, 0.00]	0.14	1.97	.048
Training	0.07 [0.00, 0.14]	0.04	1.93	.053
Spoken vs. Music \times WM	-0.02 [-0.04, 0.00]	0.01	-2.13	.033
Bg vs. Sung \times WM	0.01 [-0.03, 0.04]	0.02	0.40	.693
Sung-in vs. Sung-out \times WM	0.00 [-0.07, 0.06]	0.03	-0.03	.980
Spoken vs. Music \times Age	-0.11 [-0.34, 0.13]	0.12	-0.89	.376
Bg vs. Sung \times Age	0.03 [-0.08, 0.14]	0.06	0.56	.577
Sung-in vs. Sung-out \times Age	-0.02 [-0.09, 0.06]	0.04	-0.46	.643
Spoken vs. Music \times Training	0.01 [-0.01, 0.03]	0.01	0.78	.437
Bg vs. Sung \times Training	0.02 [-0.01, 0.05]	0.02	1.35	.176
Sung-in vs. Sung-out \times Training	0.00 [-0.07, 0.06]	0.03	-0.11	.916

Note. CI = Confidence interval; SE = Standard Error; Spoken = Spoken-in-quiet condition; Music = the three music conditions collapsed; Bg = Spoken in the presence of background music condition; Sung = the two sung conditions collapsed; Sung-in = Sung in-key condition; Sung out = Sung-out-of-key condition; WM = Working memory; Age = Age Group; Training = Years of musical training.

Cite this article: Ong, J. H. & Chan, A. H. D. (2023). Working memory modulates the effect of music on word learning *Language and Cognition* 15: 131–147. <https://doi.org/10.1017/langcog.2022.29>