

The influence of “happy” and “sad” background music on complexity, accuracy, and fluency of second-language speaking

Psychology of Music

1–15

© The Author(s) 2021



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/03057356211033345

journals.sagepub.com/home/pom

Shahaboddin Dabaghi Varnosfaderani¹ ,
Mohammadtaghi Shahnazari^{2,*}
and Azizollah Dabaghi^{2,*}

Abstract

The role of music in second-language (L2) learning has long been the object of various empirical and theoretical inquiries. However, research on whether the effect of background music (BM) on language-related task performance is facilitative or inhibitory has produced inconsistent findings. Hence, we investigated the effect of happy and sad BM on complexity, accuracy, and fluency (CAF) of L2 speaking among intermediate learners of English. A between-groups design was used, in which 60 participants were randomly assigned to three groups with two experimental groups performing an oral L2 English retelling task while listening to either happy or sad BM, and a control group performing the task with no background music. The results demonstrated the happy BM group's significant outperformance in fluency over the control group. In accuracy, the happy BM group also outdid the controls (error-free clauses, correct verb forms). Moreover, the sad BM group performed better in accuracy than the controls but in only one of its measures (correct verb forms). Furthermore, no significant difference between the groups in syntactic complexity was observed. The study, in line with the current literature on BM effects, suggests that it might have specific impacts on L2 oral production, explained by factors such as mood, arousal, neural mechanism, and the target task's properties.

Keywords

L2 speech, mood, arousal, working memory, task performance

¹Goethe University Frankfurt, Frankfurt am Main, Germany

²Department of English, Faculty of Foreign Languages, University of Isfahan, Isfahan, Iran

*The second and third authors have equal contributions in this research.

Corresponding author:

Shahaboddin Dabaghi Varnosfaderani, Goethe University Frankfurt, 60323 Frankfurt am Main, Germany.

Emails: shahab.dabaghi@stud.uni-frankfurt.de; shahabdabaghi2017@gmail.com

Students often listen to background music (BM) with or without intention while studying or performing different tasks such as reading comprehension, doing exercise, doing arithmetic operations, learning how to cook recipes, playing chess, and so on (Chew et al., 2016; De Groot & Smedinga, 2014). Previous research suggests that students who are engaged in studying may find BM either helpful and relaxing (e.g., White, 2007) or distracting (Du et al., 2020; Furnham & Bradley, 1997). The type of task students undertake and, in particular, its cognitive and linguistic characteristics may partly account for the range of the results about the impact of BM as an independent variable on task performance (De Groot & Smedinga, 2014). Research on the impact of BM on language learning and performance has primarily focused on receptive language skills, especially reading comprehension (e.g., Chew et al., 2016; Hall, 1952; Thompson et al., 2011). However, exploring the effect of BM on productive language skills such as L2 speaking appears to be notably underdeveloped (e.g., Ransdell & Gilroy, 2001).

BM as a facilitator or inhibitor

Studies on the effect of BM on language learning tasks, starting in the 1950s (e.g., Hall, 1952), have been based on a range of theoretical hypotheses and assumptions and have produced somewhat contradictory results. On one hand, much research has drawn on arousal theories such as the arousal-mood hypothesis and Yerkes–Dodson law, generally claiming that for BM to have a positive impact on learning tasks, it should evoke an appropriate mood and generate an enhanced and optimal arousal level in the task performer (Eskine et al., 2020; Thompson et al., 2011; Yerkes & Dodson, 1908). On the other hand, another group of studies has been predicated on information processing theories, such as the cognitive-capacity hypothesis and the role of working memory (WM; e.g., Salamé & Baddeley, 1989), with the concomitant assumption that there is a limited pool of cognitive resources for processing information. In this view, BM is typically seen as auditory input, potentially engaging a proportion of available cognitive resources, thus leaving less, or even insufficient, capacity for processing information related to the target task (Baddeley, 2003; Du et al., 2020; Kahneman, 1973; Lehmann et al., 2019).

Ransdell and Gilroy (2001), for instance, found that BM may reduce writing fluency. They suggested that even unattended BM could place a heavy load on participants' WM, leading to a decline in their writing fluency. However, several studies (e.g., Chew et al., 2016; Kang & Williamson, 2014; Thompson et al., 2011) pointed to a positive effect of BM in language-related task performance. They interpreted their findings based on the arousal-mood hypothesis and suggested that music enhances performance on cognitive tasks by influencing mood (i.e., typically described as having either a positive or negative affective valence, or happy vs. sad) and arousal level (e.g., high vs. low; see Husain et al., 2002). Thompson et al. (2011) argued that tempo and loudness interactively interrupt reading comprehension. As BM gets faster and louder, the adverse effect on reading comprehension, a cognitively demanding task, may increase as limited cognitive resources are diverted to processing prominent auditory events. Slow music, however, may support efficient reading as the auditory sensory system does not get excessively stimulated. Thompson et al. (2011) further suggested that a specific condition of BM might be beneficial. In their study, the scores in the fast and soft (i.e., not loud) BM condition tended to be higher than those of other conditions. They, therefore, concluded that the decisive factor in predicting whether BM helps or hinders task performance (particularly, language-related task performance) seems to be the result of a conflict in WM. This conflict takes place between the benefits of a suitable arousal level and the cost of BM's load on cognitive resources (Thompson et al., 2011).

Specific effect of music-induced moods and arousal levels

Several researchers have investigated the effect of different music-induced moods¹ and arousal levels on cognitive task performance, in particular, verbal tests. Concerning mood, for example, Bottiroli et al. (2014) demonstrated that processing of speed (in a cognitive task) was positively enhanced only by happy BM. However, in their study, semantic and episodic memories (that were, respectively, measured through free recall and phonemic fluency tests) were improved by both happy and sad BM. It is, yet, worth mentioning that the scores on the semantic memory and free recall tests were higher in the sad BM condition than those in the happy condition. Their results suggested that the assumed positive impact of happy BM, due to the positive mood creation, may not be generalizable across different types of tasks. Bottiroli et al. (2014) argued that the tasks, requiring faster processing might benefit more from happy BM, inducing a happy mood, as it could support greater mental alertness. Concerning the effect of mood, Proverbio et al. (2015) also found that emotionally touching music (e.g., sad music) as opposed to “joyful” music, could improve encoding in face recognition memory, increasing the speed of test response.

Regarding the research on the effect of arousal levels, Nguyen and Grahn (2017) also showed that low arousal music (e.g., relaxing music) could enhance scores on a recall test and low arousal music accompanied with negative mood (e.g., sad) could positively affect the performance on a recognition test; nevertheless, no improvements were observed for associative memory in all conditions (i.e., high or low arousal music with positive or negative mood). They argued that their results were in line with the Yerkes and Dodson (1908) law, suggesting that there is an “inverted U-shape” relationship between arousal level and task performance. In other words, when there is little arousal, the performance is typically poor or average. When arousal reaches a moderate level (i.e., optimal range) performance improves. However, when the arousal level exceeds the optimum level, performance decreases again. Nguyen and Grahn (2017) argued that the relaxing music in their study corresponded to moderate arousal for their participants and thus could be linked to the enhanced memory performance. Taken together, it appears plausible to assume that the mood of BM, its arousal level, and characteristics of the target task can all possibly mediate the impact of BM on cognitive and language-related task performance.

Task performance in L2 research

Information processing theories, that also form the basis of L2 task performance research (e.g., Ahmadian et al., 2015; Skehan, 2014), assume that cognitive or attentional resources for processing information are limited and, therefore, during multitasking or performing a complex task, it may not be possible to process all the information related to the task at the same level. This assumption constitutes a basic part of the trade-off hypothesis stating that allocating cognitive resources to process a specific area of linguistic performance (e.g., grammatical accuracy) is likely to come at the expense of lower performance in other areas (e.g., fluency; Ahmadian et al., 2015; Skehan, 1998; Skehan & Foster, 2001, 2014). However, the trade-off effect may possibly be compensated by manipulating the task or creating certain performance conditions (Skehan, 1998, 2014). One of these conditions to determine if the trade-off is compensated is the presence or lack of BM.

Given the research reviewed before, it could be hypothesized that BM with an optimal arousal level and cognitive load might have differential effects considering different moods such as happy and sad; and particular target tasks, for example, verbal memory, reading

comprehension, and word recall task. Therefore, the present study was carried out to address the following research question.

To what extent is performance on an L2 oral narrative production task, as measured by complexity, accuracy, and fluency (CAF), affected by sad and happy background music?

Considering the contradictory literature on BM effect on cognitive and language-related task performance, we did not make any specific predictions concerning the direction of possible effects.

Method

Design

The study employed a between-groups design with an independent variable, namely, BM with three levels (i.e., sad music, happy music, and no-music control) and a dependent variable (i.e., performance on L2 speaking task) with three levels: CAF.

Participants

Sixty Iranian undergraduate students, speaking Persian as their first language, participated in this study. They were aged between 18 and 28 years ($M = 22.23$; $SD = 3.41$), of both genders (males: $n = 32$; females: $n = 28$), and with intermediate English proficiency levels ($M = 34.46$; $SD = 4.01$). They were recruited from the total population of 150 students studying L2 English at the university level. The participants were randomly assigned to one of three groups: two experimental groups—happy BM and sad BM—and the no-music control group, with each group comprising 20 students.

The participants' L2 proficiency was determined through Version 2 of the *Oxford Placement Test* (OPT; $\alpha = .80$; Wistner et al., 2008) and their scores on an IELTS speaking mock exam (i.e., the band score 4.5 to 5.5), assessed using the *IELTS Band Descriptor* (British Council, 2018) and Common European Framework of Reference (Council of Europe, 2001). Although the OPT has been shown to be a robust test in predicting L2 English proficiency level according to its proficiency rubric, to control the confounding effect of any difference in L2 proficiency, the OPT scores of the groups were statistically compared with each other. Using a one-way analysis of variance (ANOVA; with Bonferroni post hoc analysis), no significant between-group differences between the happy versus sad versus control groups were found, $F(2, 57) = 1.05$, $MSE = 16.12$, $p = .35$.

Before the study, all of the participants expressed their informed consent to participate in the research through filling out an official consent form. The whole process of the research was permitted by the Committee of Research Ethics of the University of Isfahan and was according to the guidelines of research ethics enacted by the Ministry of Science, Research and Technology of Iran.

Materials

Music stimuli. To select the BM for the two experimental conditions, first, titles of 67 favorite instrumental songs from different music genres (e.g., jazz, hard rock, metal, classical, rap, Iranian traditional, etc.) were collected through a survey from 44 university students (age: 18–28; $M = 22.80$, $SD = 2.69$; female: $n = 28$; male: $n = 16$), asking about the main song(s) they recently listened to while doing their daily tasks. Then, a new list of songs was compiled from

the collected songs including seven songs suggested by the researchers from the literature (e.g., Nguyen & Grahn, 2017). Fifteen duplicate and inaccessible songs were also later removed from the list. Therefore, 59 songs remained on the list.

Second, 10 random participants also aged between 18 and 28 ($M = 23.90$, $SD = 2.96$; male: $n = 6$; female: $n = 4$; not the participants in the main experiment) listened to the initial 20 s of each song, rating the “happiness” versus “sadness” levels evoked by the songs. It was done through a Likert-type scale ranging from 1 to 7 as 1 referred to the highest level of *sadness* and 7 referred to the highest level of *happiness* induced by BM. The participants were also welcomed to listen to more than 20 s of each song. Then, 30 pieces of music evoking happy or sad moods, corresponding roughly to the length of time required for the oral task, were chosen according to the participants’ ratings.

Third, from this list of 30 songs, a further 10 random participants aged between 18 and 25 ($M = 20.90$, $SD = 2.80$; male: $n = 3$; female: $n = 7$) were asked to identify five happy and five sad pieces by using the updated Hevner adjective circle (Patel, 2010; Schubert, 2003). This involved listening to the first 20 s of each piece and simultaneously underlining the adjectives on the paper provided. The moods evoked by each piece were thus identified by counting the frequency of adjectives in each general category in the circle (e.g., happy, sad, angry, and frustrated). In other words, the selection of songs was based on the quantity of adjectives underlined in each category by every participant. Our main target categories in the updated Hevner’s adjective circle were Cluster A for choosing happy music and Cluster F for sad music. Cluster A contained the following adjectives: “bright,” “cheerful,” “happy,” and “joyous” and Cluster F was consisted of the following adjectives: “sad,” “dark,” “depressing,” “gloomy,” “melancholy,” “mournful,” and “solemn.” For example, for a song, if one chose four adjectives in Cluster A, three in Cluster C, and one in Cluster I, the song would be categorized under Cluster A. If in the ratings of a participant, two or more songs had similar scores in the main cluster, the scores on neighbor clusters such as B and H for ranking happy songs, and E for sad songs were considered as well.

Table A in Supplemental Material A shows the 15 songs ranked in terms of the number of participants who chose the songs based on the adjectives in Cluster A. Table B Supplemental Material A also shows the 15 songs ranked in terms of the number of participants who chose the songs based on the adjectives in Cluster F. Finally, the five most happy songs and the five most sad songs were selected based on the rankings of the songs based on the mode scores in Cluster A and Cluster F, respectively.

In addition, to rule out probable confounding effects of fatigue and habituation caused by stimuli on the ratings, the order of the songs in the playlist in each step was set as random. In this way, there is an equal chance that each song would be played in the first, middle, and last part of the music selection experiment.

Fourth, to control for the arousal level of each piece and find the two most moderately arousing songs among the happy songs and sad songs, the same participants rated the arousal level of each of the 10 selected pieces in a subsequent session, similar to the procedure in the last step, using a Likert-type scale ranging from 1 (*extremely low arousal*) to 7 (*extremely high arousal*). The *moderate arousal* level was marked as 4. Following Nguyen and Grahn (2017), the participants in this study were instructed to consider low arousal as “relaxing” and “calming” and high arousal as “distressing” and “thrilling.” Moreover, to avoid participants assigning subjectively different numbers to an arousal level, each number in the scale was marked with a certain arousal level. This would help to minimize heterogeneity in data.

Subsequently, the number 4 (i.e., the *moderate arousal* on the scale) was considered as a representation for moderate arousal level evoked by a song, and based on their choices, the two

most selected songs (i.e., one happy and one sad) in terms of moderate arousal level for the speaking task performance were chosen. In other words, the songs with the greatest mode score of *moderate arousal* were selected.

As indicated in Table C in Supplemental Material A, among the songs, “The Rap” (*Mode* = 4, *M* = 4.2, *SD* = 0.78) and “Song from a Secret Garden” (*Mode* = 4, *M* = 4.0, *SD* = 0.47) had the highest mode score in the number 4 and the closest mean score to 4. Therefore, the final selected musical stimuli were “The Rap” as the instrumental happy BM and “Song from a Secret Garden” as the instrumental sad BM, both composed by the group known as Secret Garden (Løvland, 1995). The pieces were 2 min 31 s and 3 min 32 s long, respectively, but each piece was set to replay when it ended in case the participants continued to talk. The arousal level ratings of the final songs were also compared with each other through a Chi-square test of independence to check for any potential difference, $\chi^2(df = 6, N = 10) = 10.35$, $p = .11$, Cramer’s $V = .72$.

Given that there were significant dynamics in terms of volume within the pieces, they were manipulated using Cubase 5 software to ensure constant volume throughout. In addition, an appropriate volume of the music (i.e., ~60 dB) was set by the same participants through Cubase 5 software. Moreover, to rule out probable confounding effects of fatigue and habituation caused by stimuli on the ratings, the order of the songs in the playlist in each step was set as random. In this way, there is an equal chance that each song would be played in the first, middle, and last part of each stage of the music selection experiment. All of the participants in this study were Iranian and Persian was their only first language.

Speaking task. Following Mackey and Gass (2005), an oral retelling task was developed, based on a written text followed by three prompt questions. According to Varnosfadrani (2006), the retelling task was based on a written narrative entitled “Britain’s Unluckiest Criminal,” taken from *Cutting Edge* by Carr et al. (2001). The three task questions were designed to ask the participant to retell the story they had just read, to tell the researcher the reason why the main character of the story wanted to commit suicide, and what would the participant do if they were the main character of the story (and why?) (see Supplemental Material E).

The reading difficulty as determined by the Fry Readability Graph was “fairly easy to read” and “suitable” for L2 students with intermediate English proficiency (Chinn et al., 2014). The task’s questions were as follows: 1. “Please tell the story that you have just read.” 2. “In your opinion, why do you think Carson (i.e., the main character of the story) wanted to kill himself?” 3. “If you were in his position, what would you do instead? What are the reasons for your actions?” Further, the task was considered to be at the B1 level (i.e., intermediate level) of the Common European Framework of Reference (Council of Europe, 2001).

Procedure

The 60 participants completed the reading and oral retelling task individually in succession in a quiet room using an Acer Aspire E 15 laptop. Each participant, in turn, was instructed to click on a preloaded PowerPoint presentation which first gave instructions about the task in their L1, followed by a page providing definitions and glosses of key terms used in the text such as “roulette” and “cliff” which might be opaque linguistically or culturally, followed by the reading text. Four minutes were allowed for reading the text. The slide then changed to present the first question. The specific BM was automatically played to those in the experimental groups through headphones, starting from the time the participants began reading the first question and answering it orally, and lasting until they had completed the oral task by

answering all the questions. Participants' answers to the questions were recorded using a Sony HD recorder, activated throughout the question-and-answer part of the task ($M = 3$ min 51 s, $SD = 1.86$). The whole test lasted about ~ 10 min. All instructions, written or spoken, given in the experiment were double-checked by the researchers and by an L2 expert to avoid potential misinterpretation.

Data analysis

Scoring procedure for CAF. First, the recordings of participants' responses were transcribed using Dragon Naturally Speaking software. Based on the procedures used in Ahmadian et al. (2015) and Sato and Lyster (2012), outlined below, indices of CAF in the transcriptions were counted by the first author. Finally, the transcriptions and the scorings were checked by another rater (independent of the study).

To measure syntactic complexity, the number of the clauses, spoken by each of the participants, was divided by the number of AS-units (i.e., the Analysis of Speech unit) in their recorded speech (Ahmadian et al., 2015). A meaningful unit of oral speech consists of a main clause plus any relevant subordinate clause(s). It represents an independent mental unit of oral linguistic processing that is regarded as an AS-unit. However, in considering this unit of oral speech, certain factors such as long pauses, repetitions, and changes in ideas as the unit barrier markers, are also taken into account (for a discussion see Foster et al., 2000).

Sato and Lyster (2012), following Foster and Skehan (1996), define "accuracy" as language production which is free from morphological, syntactic, and word order errors, measured by the ratio of error-free clauses to the total clauses in the speech. The percentage of correct verb forms (i.e., subject-verb agreement, and correct use of aspect, modality, and tense) was also calculated as a measure of speech accuracy (Sato & Lyster, 2012, as cited in Ahmadian et al., 2015).

To measure fluency, we followed Ahmadian et al. (2015) in using two indices: Rate A and Rate B. Rate A is based on the number of syllables within each narrative divided by the number of seconds used to complete the task and then multiplied by 60. Rate B is based on the number of meaningful syllables produced per minute, thus a similar measurement to Rate A but excluding repeated, reformulated, or replaced words.

Reliability of the scorings. To determine the reliability of the CAF scoring, a sample of 25% of the dataset was checked by another rater. The mean of interrater reliability for all indices was 0.92.

Statistical procedures. Following Pallant (2011), the data were checked against normality tests to choose either a parametric or nonparametric test for each measure of the dependent variables. Although the participants in this study were randomly assigned to the three groups, the results of the Shapiro-Wilk test for normality (see Supplemental Material B) showed that the accuracy scores (e.g., correct verb forms) were significantly non-normal ($p < .001$). Moreover, homogeneity of variance(s), dependent on meeting normal distribution, was checked before considering parametric tests. All of the measures had sufficient homogeneity of variance, except Rate B. Subsequently, Rate B was checked against the robust tests of equality of means (e.g., Brown-Forsythe test) and violated them. Therefore, to compare Rate B, correct verb forms, and error-free clauses between the groups (i.e., sad BM group vs. happy BM group vs. control group), nonparametric tests (i.e., Kruskal-Wallis Test for one-way between-groups analysis of variance with post hoc pairwise analysis using the Bonferroni significance level adjustment) were used.

Table 1. Descriptive Statistics of Dependent Variable Measures Among the Control, Happy, and Sad Conditions.

Dependent variables' measures	All groups	<i>M</i>	<i>Mdn</i>	<i>SD</i>	95% CI for <i>M</i>	
					LL	UL
Syntactic complexity	Control	139.44	142.02	14.93	132.45	146.43
	Happy	146.77	145.34	14.62	139.93	153.62
	Sad	149.16	150.00	19.66	139.96	158.36
Error-free clauses	Control	61.56	58.11	18.13	53.07	70.05
	Happy	76.31	82.08	17.01	68.35	84.27
	Sad	66.79	70.29	21.97	56.50	77.07
Correct verb forms	Control	61.49	61.59	21.10	51.61	71.37
	Happy	83.91	87.93	14.16	77.28	90.54
	Sad	77.99	87.94	20.27	68.50	87.47
Rate A	Control	89.37	79.10	25.28	77.54	101.20
	Happy	119.20	117.35	27.42	106.37	132.04
	Sad	108.05	111.02	38.91	89.83	126.26
Rate B	Control	79.90	74.78	26.18	67.65	92.16
	Happy	109.85	107.72	24.18	98.53	121.17
	Sad	101.13	103.80	40.37	82.24	120.03

SD: standard deviation; *CI*: confidence interval; *LL*: lower limit; *UL*: upper limit.

Rate A also violated the normality test; however, parametric procedures such as ANOVA have been reported to be fairly robust against slight violations of normality (Pallant, 2011). Thus, Rate A and syntactic complexity were analyzed through 2 one-way ANOVAs with post hoc analysis using the Bonferroni significance level adjustment. All of the statistical procedures were undertaken using the SPSS 24 software.

Results

Table 1 provides the descriptive statistics in detail and Figure 1 depicts the higher mean scores of the sad and happy BM conditions compared to the control condition. Table 2, giving the results of the ANOVA test, shows that there was a significant difference between all the groups in Rate A. However, no significant difference in syntactic complexity between the groups was observed. Moreover, as shown in Table 3, a Kruskal–Wallis test demonstrated the significant differences between all the groups in error-free clauses, correct verb forms, and Rate B with higher medians and mean scores of happy and sad BM groups (see Table 1).

To understand which layers of the independent variable (i.e., sad or happy BM) affected the measures of CAF, after the ANOVA and Kruskal–Wallis tests, multiple post hoc pairwise analyses with Bonferroni significance level adjustment were conducted. The results showed that, in Rate A, Rate B, and error-free clauses, only the happy BM condition had significantly higher scores than the scores of the control condition (see Supplemental Materials C and D). However, in correct verb forms, both happy and sad BM groups significantly outperformed the controls. Additionally, no significant pairwise differences between the happy and sad groups were observed.

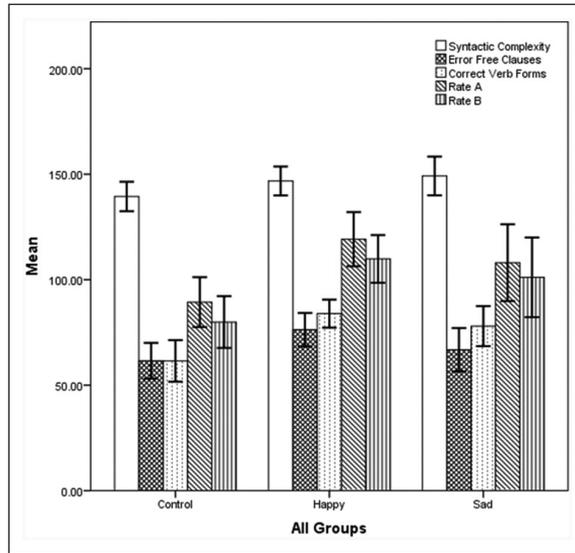


Figure 1. Comparison of the Measures of Dependent Variables Between the Sad, Happy, and Control Conditions Presented With Standard Error Bars (With 95% Confidence Interval).

Table 2. Overall Differences between the Sad BM, Happy BM, and Control Conditions in Syntactic Complexity and Rate A.^a

Dependent variables' measures	Source	Sum of squares	Mean square	F(2, 57)	p	η ²
Syntactic complexity	Between groups	1,025.56	512.78	1.86	.16	.06
	Within groups	15,650.68	274.57			
	Total	16,676.24				
Rate A	Between groups	9,089.24	4,544.62	4.69	.01*	.14
	Within groups	5,5213.53	968.65			
	Total	64,302.78				

^aOne-way between-groups analysis of variance (ANOVA).

*The mean difference is significant at the .05 level.

Table 3. Overall Differences Between the Sad BM, Happy BM, and Control Conditions in Error-Free Clauses, Correct Verb Forms, and Rate B.^a

	Error-free clauses	Correct verb forms	Rate B
Test statistic	6.38	11.32	8.50
df	2	2	2
Asymp. Sig.	.04*	.00**	.01*
Power (1 – β err prob)	.53	.87	.73
Eta squared (η ²)	.07	.16	.11
Cohen's d	.57	.88	.71

^aKruskal–Wallis test.

*p < .05. **p < .01.

Discussion

This study was done to investigate whether instrumental BM, in particular, happy or sad BM, influences intermediate L2 learners' speaking performance on an oral retelling task, as measured through CAF. The results indicate that happy BM has significant impacts on accuracy and fluency, but sad BM's positive effect may be restricted to correct verb forms as one of the measures of L2 accuracy. Nevertheless, both BM conditions do not affect syntactic complexity and, overall, do not differ significantly from each other in CAF when compared in pairs.

Effect of the BM on fluency and accuracy

The findings of the present study are mostly consistent with those in previous studies such as Thompson et al. (2011) and Kang and Williamson (2014). They could be generally discussed in terms of the arousal-mood hypothesis and the Yerkes–Dudson law (Thompson et al., 2011; Yerkes & Dodson, 1908). These two theories propose a positive effect of appropriate arousal level and specific mood evoked by BM on language-related task performance (Thompson et al., 2011). In other words, BM might have provided a better performing context by modulating emotional excitation and keeping the arousal level in an optimal range.

Despite the explanation considering the arousal-mood hypothesis, the better performance of the happy group could also be justified by the enhanced mental alertness triggered by upbeat music (Bottiroli et al., 2014). They, for example, found positive effects of happy music on phonemic fluency and speed processing. The authors hypothesized that tasks which demand speed in processing (e.g., fluency in speaking) require a more alert state of mind. Hence, happy and major scale music, which could trigger a more alert mood in the listener, might be more beneficial than sad music (i.e., mostly characterized by a minor scale) in supporting oral fluency. These findings are also in line with those of previous related studies (e.g., Freggens, 2014; Nguyen & Grahn, 2017; Proverbio et al., 2015), suggesting specific effects for moods evoked by BM.

Another account for the better performance of the happy group is that happy and major scale BM is reported to be beneficial to memory function (Ferreri et al., 2014; Mammarella et al., 2007). The neuroscientific literature (e.g., Blumenfeld & Ranganath, 2007) about patients with a deficiency in the prefrontal cortex, showing its responsibility for encoding verbal information and WM—particularly, the episodic buffer—revealed that BM could reduce the activation load of the cortex (Ferreri et al., 2014). It could, therefore, facilitate the processing of complex cognitive tasks such as speaking. In fact, the cortex supports the creation of connections between active pieces of information in WM and a number of memory encoding processes involved in linguistic processing, for example, mental organization, word association, and meaning.

Thus, BM, which decreases the neural burden of the information on the cortex, could provide a better situation for the encoding of the information related to the processes (Ferreri et al., 2014). To elaborate, the limbic-reward loop in the brain is reported to be activated before memory formation (Adcock et al., 2006), and BM is shown to be a significant rewarding stimulus in subcortical and limbic-reward brain regions (Salimpoor et al., 2013). For this reason, a link between the lower neural activation load of the prefrontal lobe and the prior activation of the limbic-reward loop under BM condition could exist, implying the compensatory role of the limbic-reward system in lowering activation of the prefrontal cortex which could explain the better performance of the cortex and BM effect on verbal learning, memory, and performance (Ferreri et al., 2014).

Effect of the BM on syntactic complexity

Regarding complexity, however, as Ellis (2003) suggested, oral production of syntactic complexity can be a highly demanding task considering the limitations in attentional resources of WM. Hence, it could be cognitively more challenging for intermediate L2 learners to process complex structures (i.e., independent and subordinate clauses) when listening to BM. Furthermore, syntactic complexity might not be solely dependent on the efficiency of WM in retrieving linguistic knowledge and it might involve other factors such as personality traits and self-confidence in using recently acquired complex grammar in speaking (Ellis, 2003). This assumption is supported by the findings of previous studies such as Tavares (2008) and Ahmadian (2012), reporting no significant relationship between L2 complexity and WM capacity. Relevantly, Anderson (1983) argues that complex structures are gradually internalized in language learners' memory as implicit knowledge—as a result of *proceduralization*. This implies that complex syntactic structures require extensive practice for development and, thus, might be more affected by the performer's implicit knowledge of the structures rather than the limitations of attentional resources in WM.

Pairwise comparison between happy and sad BM groups

Interestingly, partly in contrast to Nguyen and Grahn's (2017) results, no significant pairwise differences between the happy and sad BM groups were observed. One account for this could be a difference in methodology of this study and theirs: In this study, musical stimuli were selected within an optimal arousal range, but they were chosen based on closeness to the middle number of the measuring Likert-type scale (i.e., the number 4 that was regarded as the *moderate arousal*). However, Nguyen and Grahn (2017), taking both Yerkes–Dodson law and arousal-mood hypothesis into account, rather chose their musical stimuli by considering a high and low arousal cut-off criterion in their scale. It means that their songs had either low or high arousal levels but within an optimal arousal range (to avoid overstimulation in selecting their musical stimuli) similar to one in this study. One implication of this is to hypothesize that certain arousal levels within an optimal range could possibly have particular direct or interactive effects on task performance.

Thus, considering the studies such as Nguyen and Grahn (2017) that proposed interaction between mood and arousal for BM effect, the lack of significant differences between the sad and happy BM conditions in this study might be attributed to similar arousal levels of the songs chosen for this study. This might probably have reduced the music-induced mood effects so that no significant pairwise differences between them were observed.

Conclusion, limitations, and implications of the study

In line with the arousal-mood hypothesis (Thompson et al., 2011; Yerkes & Dodson, 1908) and the relevant neurocognitive research (Ferreri et al., 2014; Mammarella et al., 2007), this study shows a general positive effect of BM on L2 speaking performance measured through CAF among intermediate learners of English in a single session. In addition, based on the literature (Blumenfeld & Ranganath, 2007; Ferreri et al., 2014; Mammarella et al., 2007) showing the impact of upbeat music on reducing the burden on the prefrontal cortex, this study could demonstrate the positive specific effect of the happy BM on L2 oral accuracy and fluency, and the sad BM on the quantity of correct verb forms as an index of L2 oral accuracy. By showing the non-significant results of comparing happy and sad conditions in pairs, this research also

demonstrates how music-induced mood effects could possibly have a much intricate mechanism and/or interactive relationship with factors such as arousal level.

The present study also has its own limitations. Although this research may have enjoyed random distribution in group assignment and no data or sample attrition, it partly suffered from data non-normality. This might stem from the nature of the implemented measures, dependent variables, or unknown intervening variables. Moreover, this research would have more assertive implications with further intervention sessions and repeated measures to see if and how long such results can last. Therefore, more precise random sampling of a larger number of participants with additional treatment sessions and delayed tests, which was beyond the scope of this study, might better deduce the effect of unknown variables on data normality and offer stronger implications.

Barring the mentioned constraints, the study may have implications for L2 teachers, L2 theorists, and music psychology experts. It opens a perspective in research on how, for instance, BM can have a positive effect on learning a specific L2 feature or performance on an L2 task in a realistic longer individual or classroom trial. Furthermore, to the best of the researchers' knowledge, the sole aspect of this study is that, to date, no attempt has been made to explore the effect of BM among L2 learners with an intermediate level of proficiency in speaking. Moreover, no efforts have been made to investigate BM effect in view of individual differences such as WM capacity among these learners. Therefore, the present study can make a contribution to the field, giving rise to revisiting emotional factors in learning and teaching a second language. It can provide implications for research on the influence of musically created moods in L2 learning in light of the studies on the role of emotions in L2 learning and teaching (Dewaele et al., 2019; Gregersen, 2016; Shao et al., 2019).

Additionally, it is a fruitful suggestion to triangulate a behavioral study like the present one, using neuroscientific methods (e.g., EEG and fMRI) which may provide the researchers in the field with neural evidence (Du et al., 2020). They might shed more light on the mechanisms and processes underlying the possible influences of BM by monitoring, for example, the physiological arousal level of the stimulus and the activated neural network. They might also pave the way for understanding the potential of the relationship between the use of music and L2 learning. Furthermore, a new replication of this study in terms of different trending theoretical perspectives such as the seductive details hypothesis (Lehmann et al., 2019) is encouraged. According to Beer and Greitemeyer (2019), BM has also been shown to be unavoidably impactful in social behavior. Thus, investigating verbal social behavior (in the first language and/or L2) in relation to BM can be another relevant and socially implicative field of inquiry.

Acknowledgement

We would like to express our appreciation to the respectful reviewers whose constructive comments immensely enhanced the manuscript. Special thanks to Prof. Mohammad Javad Ahmadian and Dr. Britta Biedermann for their helpful reviews on the early versions of the manuscript, Prof. Daniela Elsner for her academic and moral support, and Mr. Mehdi Saraeian and Peiman Sattari for their technical help.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Shahaboddin Dabaghi Varnosfaderani  <https://orcid.org/0000-0001-7873-5013>

Supplemental material

Supplemental material for this article is available online.

Note

1. The term *Mood* in this research refers to a short living psychological response or state induced by one-time playing of a song as a stimulus (Thayer, 1989). We used the word mood because it may mostly assimilate to the term “mood,” used in the arousal-mood hypothesis and used by Nguyen and Grahn (2017). However, for a detailed discussion on the theoretical and operational definitions of mood and emotion in psychology, please refer to Beedie et al. (2005).

References

- Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. D. (2006). Reward-motivated learning: Mesolimbic activation precedes memory formation. *Neuron*, *50*(3), 507–517. <https://doi.org/10.1016/j.neuron.2006.03.036>
- Ahmadian, M. J. (2012). The relationship between working memory capacity and L2 oral performance under task-based careful online planning condition. *TESOL Quarterly*, *46*(1), 165–175. <https://doi.org/10.1007/s10936-013-9253-7>
- Ahmadian, M. J., Tavakoli, M., & Dastjerdi, H. (2015). The combined effects of online planning and task structure on complexity, accuracy, and fluency of L2 speech. *Language Learning Journal*, *43*(1), 41–56. <https://doi.org/10.1080/09571736.2012.681795>
- Anderson, J. R. (1983). *The architecture of cognition*. Harvard University Press. [https://doi.org/10.1016/0004-3702\(86\)90084-6](https://doi.org/10.1016/0004-3702(86)90084-6)
- Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*, 829–839. <https://doi.org/10.1038/nrn1201>
- Beedie, C., Terry, P., & Lane, A. (2005). Distinctions between emotion and mood. *Cognition and Emotion*, *19*(6), 847–878. <https://doi.org/10.1080/02699930541000057>
- Beer, A., & Greitemeyer, T. (2019). The effects of background music on tipping behavior in a restaurant: A field study. *Psychology of Music*, *47*(3), 444–450. <https://doi.org/10.1177/0305735618755887>
- Blumenfeld, R. S., & Ranganath, C. (2007). Prefrontal cortex and long-term memory encoding: An integrative review of findings from neuropsychology and neuroimaging. *The Neuroscientist*, *13*(3), 280–291. <https://doi.org/10.1177/1073858407299290>
- 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*, Article 284. <https://doi.org/10.3389/fnagi.2014.00284>
- British Council. (2018). *IELTS band descriptors*. <https://takeielts.britishcouncil.org/teach-ielts/test-information/assessment>
- Carr, J. C., Cunningham, S., & Moor, P. (2001). *Cutting edge: Pre intermediate teacher's resource book*. Longman.
- Chew, A. S. Q., Yu, Y. T., Chua, S. W., & Gan, S. K. E. (2016). The effects of familiarity and language of background music on working memory and language tasks in Singapore. *Psychology of Music*, *44*(6), 1431–1438. <https://doi.org/10.1177/0305735616636209>
- Chinn, L., McGuirt, S., & Puri, S. (2014). The readability of online patient education materials from major anesthesiology associations and the American society of anesthesiologists. *Open Journal of Anesthesiology*, *4*(1), Article 41001. <https://doi.org/10.4236/ojanes.2014.41001>
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press. <https://www.cambridge.org/de/cambridgeenglish/catalog/teacher-training-development-and-research/common-european-framework-reference-languages/common-european-framework-reference-languages-learning-teaching-assessment-paperback?isbn=9780521005319&format=PB>

- De Groot, A. M., & Smedinga, H. E. (2014). Let the music play!. *Studies in Second Language Acquisition*, 36(04), 681–707. <https://doi.org/10.1017/S0272263114000059>
- Dewaele, J. M., Chen, X., Padilla, A. M., & Lake, J. (2019). The flowering of positive psychology in foreign/second language teaching and acquisition research. *Frontiers in Psychology*, 10, Article 2128. <https://doi.org/10.3389/fpsyg.2019.02128>
- Du, M., Jiang, J., Li, Z., Man, D., & Jiang, C. (2020). The effects of background music on neural responses during reading comprehension. *Scientific Reports*, 10(1), 1–10. <https://doi.org/10.1038/S41598-020-75623-3>
- Ellis, R. (2003). *Task-based language teaching and learning*. Oxford University Press. <https://global.oup.com/academic/product/task-based-language-learning-and-teaching-9780194421591?lang=3n&cc=fi>
- Eskine, K. E., Anderson, A. E., Sullivan, M., & Golob, E. J. (2020). Effects of music listening on creative cognition and semantic memory retrieval. *Psychology of Music*, 48(4), 513–528. <https://doi.org/10.1177/0305735618810792>
- 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, Article 301. <https://doi.org/10.3389/fnhum.2014.00301>
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18, 293–323. <https://doi.org/10.1017/S0272263100015047>
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21(3), 354–375. <https://doi.org/10.1093/applin/21.3.354>
- Freggens, M. (2014). *The effect of music type on emotion regulation: An emotional-Stroop experiment* [Honors Thesis]. Georgia State University. https://scholarworks.gsu.edu/psych_hontheses/19
- Furnham, A., & Bradley, A. (1997). Music while you work: The differential distraction of background music on the cognitive test performance of introverts and extraverts. *Applied Cognitive Psychology*, 11(5), 445–455. [https://psycnet.apa.org/doi/10.1002/\(SICI\)1099-0720\(199710\)11:5%3C445::AID-ACP472%3E3.0.CO;2-R](https://psycnet.apa.org/doi/10.1002/(SICI)1099-0720(199710)11:5%3C445::AID-ACP472%3E3.0.CO;2-R)
- Gregersen, T. (2016). The positive broadening power of a focus on well-being in the language classroom. In D. Gabryś-Barker & D. Galajda (Eds.), *Positive psychology perspectives on foreign language learning and teaching: Second language learning and teaching* (pp. 59–73). Springer. https://doi.org/10.1007/978-3-319-32954-3_4
- Hall, J. C. (1952). The effect of background music on the reading comprehension of 278 eighth and ninth grade students. *Journal of Educational Research*, 45(6), 451–458. <https://doi.org/10.1080/00220671.1952.10881962>
- Husain, G., Thompson, W. F., & Schellenberg, E. G. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music Perception: An Interdisciplinary Journal*, 20(2), 151–171. <https://doi.org/10.1525/mp.2002.20.2.151>
- Kahneman, D. (1973). *Attention and effort*. Prentice-Hall. https://scholar.princeton.edu/sites/default/files/kahneman/files/attention_hi_quality.pdf
- Kang, H. J., & Williamson, V. J. (2014). Background music can aid second language learning. *Psychology of Music*, 42(5), 728–747. <https://doi.org/10.1177/0305735613485152>
- Lehmann, J. A. M., Hamm, V., & Seufert, T. (2019). The influence of background music on learners with varying extraversion: Seductive detail or beneficial effect? *Applied Cognitive Psychology*, 33, 85–94. <https://doi.org/10.1002/acp.3509>
- Løvland, R. (1995). *Songs from a secret garden* [Album recorded by Secret Garden]. Universal Music Norway.
- Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. Lawrence Erlbaum Associates. <https://www.routledge.com/Second-Language-Research-Methodology-and-Design/Mackey-Gass/p/book/9781315750606>
- Mammarella, N., Fairfield, B., & Cornoldi, C. (2007). Does music enhance cognitive performance in healthy older adults? The Vivaldi effect. *Aging Clinical and Experimental Research*, 19(5), 394–399. <https://doi.org/10.1007/BF03324720>
- 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>

- Pallant, J. (2011). *Survival manual: A step by step guide to data analysis using SPSS*, 4. Allen & Unwin. <http://spss4.allenandunwin.com.s3-website-ap-southeast-2.amazonaws.com/default.html>
- Patel, A. D. (2010). *Music, language, and the brain*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195123753.001.0001>
- Proverbio, A. M., Nasi, V. L., Arcari, L. A., De Benedetto, F., Guardamagna, M., Gazzola, M., & Zani, A. (2015). The effect of background music on episodic memory and autonomic responses: Listening to emotionally touching music enhances facial memory capacity. *Scientific Reports*, 5, Article 15219. <https://doi.org/10.1038/srep15219>
- Ransdell, S. E., & Gilroy, L. (2001). The effects of background music on word processed writing. *Computers in Human Behavior*, 17(2), 141–148. [https://doi.org/10.1016/S0747-5632\(00\)00043-1](https://doi.org/10.1016/S0747-5632(00)00043-1)
- Salamé, P., & Baddeley, A. (1989). Effects of background music on phonological short-term memory. *The Quarterly Journal of Experimental Psychology*, 41(1), 107–122. <https://psycnet.apa.org/doi/10.1080/14640748908402355>
- Salimpoor, V. N., van den Bosch, I., Kovacevic, N., McIntosh, A. R., Dagher, A., & Zatorre, R. J. (2013). Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science*, 340(6129), 216–219. <https://doi.org/10.1126/science.1231059>
- Sato, M., & Lyster, R. (2012). Peer interaction and corrective feedback for accuracy and fluency development: Monitoring, practice, and proceduralization. *Studies in Second Language Acquisition*, 34(4), 591–626. <https://doi.org/10.1017/S0272263112000356>
- Schubert, E. (2003). Update of the Hevner adjective checklist. *Perceptual and Motor Skills*, 96(3_suppl), 1117–1122. <https://doi.org/10.2466/pms.2003.96.3c.1117>
- Shao, K., Pekrun, R., & Nicholson, L. (2019). Emotions in classroom language learning: What has been done and what can be learned. *System*, 86, Article 102121. <https://doi.org/10.1016/j.system.2019.102121>
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press. <https://doi.org/10.1093/elt/54.4.398>
- Skehan, P. (Ed.) (2014). *Processing perspectives on task performance* (Vol. 5). John Benjamins. <https://doi.org/10.1075/tblt.5>
- Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 83–205). Cambridge University Press. <https://doi.org/10.1017/CBO9781139524780>
- Tavares, M. D. G. G. (2008). *Pre-task planning, working memory capacity, and L2 speech performance* [Doctoral dissertation]. Federal University of Santa Catarina. <http://www.repositorio.ufc.br/handle/riufc/19548>
- Thayer, R. E. (1989). *The biopsychology of mood and arousal*. Oxford University Press. <https://global.oup.com/academic/product/biopsychology-of-mood-and-arousal-9780195068276?cc=de&lang=en>
- Thompson, W. F., Schellenberg, E. G., & Letnic, A. K. (2011). Fast and loud background music disrupts reading comprehension. *Psychology of Music*, 40(6), 700–708. <https://doi.org/10.1177/02F0305735611400173>
- Varnosfadrani, A. D. (2006). *A comparison of the effect of implicit/explicit and immediate/delayed corrective feedback on learners' performance in Tailor-made tests* [Doctoral dissertation]. The University of Auckland. <https://researchspace.auckland.ac.nz/handle/2292/55>
- White, K. N. (2007). The effects of background music in the classroom on the productivity, motivation, and behavior of fourth grade students [Unpublished master's thesis]. *Columbia College*. <https://files.eric.ed.gov/fulltext/ED522618.pdf>
- Wistner, B., Sakai, H., & Abe, M. (2008). An analysis of the Oxford placement test and the Michigan English placement test as L2 proficiency tests. *Journal of Takasaki City*, 125(50), Article 134. <https://core.ac.uk/download/pdf/223191506.pdf>
- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurology and Psychology*, 18, 459–482. <https://doi.org/10.1002/cne.920180503>