

Research paper

Effect of Songs and Stories on Elementary Male and Female EFL Learners'

Vocabulary Learning/Retention

Behnaz Alijani*¹, Mahmood Hashemian², Aliakbar Jafarpour³

^{1,2,3} Department of English Language, Faculty of Letters & Humanities, Shahrekord University, Iran

Alijani, B., Hashemian, M., & Jafarpour, A. (2026). Effect of songs and stories on elementary male and female EFL learners' vocabulary learning/retention. *Journal of new advances in English Language Teaching and Applied Linguistics*, 8(1), 174-199.

Doi:10.22034/Jeltal.2026.8.1.8

* Corresponding Author: behnaz.alijani74@gmail.com

Abstract

This study examined the (possible) impact of songs and stories as incidental ways on elementary (fe)male EFL learners' vocabulary learning/retention. To this end, 60 elementary (fe)male EFL students, within the age range of 8-12, in 2 language institutes in Isfahan province, Iran, were chosen. They were assigned to 4 groups: 2 song groups, each with 15 (fe)male participants; and 2 story groups, each with 15 (fe)male participants. All learners were at the same proficiency level, determined by the Oxford Quick Placement Test (OQPT). To collect the data, vocabulary tests with 20 items (fill-in-the-blanks, multiple-choice, true/false, and matching) were administered as pretest, posttest, and delayed posttest. The data were analyzed using one-way repeated measures ANOVA and two-way ANCOVA. Findings indicated that both songs and stories significantly improved participants' L2 vocabulary knowledge, but the song groups outperformed the story groups in both learning and retention. Examination of gender differences showed no significant differences in vocabulary learning between male and female participants. In terms of retention, females in both groups slightly outperformed males, though the differences were not considerable. These findings suggest that incorporating songs and stories into EFL instruction can enrich vocabulary knowledge for young learners.

Keywords: EFL learners, Songs, Stories, Vocabulary learning/retention

Received: 2025-08-22 **Revised** 2025-12-01 **Accepted** 2026-03-08

Introduction

Vocabulary has long been regarded as one of the most essential components of L2 proficiency (Carter & McCarthy, 2014). As Abdollahzadeh and Amiri-Vardani (2010) argue,

vocabulary knowledge forms the core of language competence because, without it, learners cannot effectively comprehend or produce the target language. Schmitt (2008) likewise emphasizes that a broad and deep lexical repertoire is crucial for successful L2 development. Similarly, Qian (2002) asserts that vocabulary knowledge plays a central role in supporting other language skills, particularly reading comprehension. Recent research continues to underscore these claims, demonstrating that vocabulary knowledge is one of the strongest predictors of overall language achievement in EFL contexts (Alharthi, 2025).

In the context of promoting lexical acquisition, two pedagogical approaches have been widely recognized: explicit and incidental instruction. Explicit instruction involves deliberate and structured teaching of vocabulary, whereas incidental learning occurs through exposure to words embedded in meaningful, context-rich input (Penno et al., 2002). Over the past decades, incidental vocabulary acquisition has attracted considerable scholarly interest, especially through modalities such as reading (e.g., Horst, 2005; Webb, 2007) and listening (e.g., van Zeeland & Schmitt, 2013; Vidal, 2011). A recent meta-analysis by Webb et al. (2023) confirms that incidental vocabulary learning through reading, listening, and reading-while-listening results in substantial vocabulary gains, highlighting the sustained effectiveness of meaning-focused exposure. Within this framework, songs and stories have been recognized as effective sources of incidental vocabulary development, particularly for children and elementary-level EFL learners.

The use of songs and music represents a promising strategy for enhancing L2 learners' motivation and engagement in vocabulary learning. As Schön et al. (2007) state, "songs are ubiquitous" (p. 2) and serve multiple functions. According to Trainor, Austin, and Desjardins (2000), the most prominent function of songs is their affective appeal. Beyond emotional benefits, songs also support cognitive and linguistic processes (Schoepp, 2001). For example, Medina (1993) demonstrated that songs significantly facilitate vocabulary learning in young L2

learners, a finding supported by more recent research showing that adolescent ESL learners who sang songs with familiar melodies exhibited improved vocabulary recall and pronunciation (Zhang et al., 2023). Additional studies (Ayotte, 2004; Fischler, 2006; Neumann, 2004; Ransdell & Gilroy, 2001; Schön et al., 2004) further highlight the positive effects of integrating songs into L2 instruction. McLaren and Lankshear (1994) argue that using songs in L2 teaching can support the development of phonological, grammatical, lexical, cultural, and sociolinguistic competencies, regardless of musical genre, training, or instructional style. More broadly, songs can contribute meaningfully to the development of learners' listening, speaking, reading, and writing skills while providing rich opportunities to acquire vocabulary, grammatical structures, pronunciation patterns, and rhythmic features of the language (Lo & Li, 1998).

Stories and storytelling techniques also hold significant value in educational settings worldwide. Soleimani and Akbari (2013) emphasize that stories serve as rich resources for children, enabling them to gain insights into the world around them, as real-life experiences are vividly reflected within the imaginative space of storytelling. As Celce-Murcia (2001) asserts, many L2 teachers of young learners carefully select stories from children's literature because they offer engaging materials that support learners psychologically, socially, culturally, linguistically, and cognitively. Brown (2007) similarly notes that teachers incorporate stories either as supplements to their main instructional materials or as stand-alone units that function as mini-syllabi. Thus, stories provide a meaningful foundation and a rich context for fostering a range of interconnected language-learning tasks that actively and creatively engage young learners within a holistic curriculum (Ellis & Brewster, 2002). Drawing on the natural approach, Krashen (1981) maintains that children are natural L2 learners who acquire language more effectively than adults when immersed in authentic, enjoyable, and meaningful learning environments.

Nyikos and Fan (2007) identified several factors that influence second language vocabulary acquisition, including individual differences, gender, proficiency level, learning context, and learning strategies. However, as noted by Pourshahian et al. (2012), the role of gender in L2 vocabulary development remains relatively underexplored. Therefore, the present study examines English vocabulary learning through stories and songs and explores potential gender differences in learners' vocabulary acquisition and retention.

Aligned with the objectives of the current study, the following research questions were formulated:

1. Does contact with L2 input via songs have any significant impact on elementary EFL students' learning/retention of lexicon?
2. Does contact with L2 input via stories have any significant impact on elementary EFL students' learning/retention of lexicon?
3. Does the gender of elementary EFL students have any significant moderating role in vocabulary learning/retention via songs and stories?

Literature Review

Scholars such as Heidari (2019), Qian and Lin (2020), and Read (2000) have emphasized the centrality of vocabulary acquisition in effective second language use and highlighted its essential contribution to coherent spoken and written communication. Supporting this view, Nation (2022) asserted that vocabulary plays a fundamental role across all language skills—speaking, reading, listening, and writing—in both ESL and EFL contexts. Likewise, acquiring a substantial lexical repertoire is necessary for successful L2 functioning because, without extensive vocabulary knowledge, learners are unable to employ the structures and functions they have learned to communicate comprehensibly (Rivers & Nunan, 1991).

Songs, like other pedagogical tools (e.g., flashcards, pictures, games), are considered an effective means for teaching vocabulary, particularly to young L2 learners. Medina (1993) noted that songs enable learners to acquire a variety of useful lexical items. Murphey (1992a) argued that large sets of vocabulary can be taught through the repetition inherent in popular songs. Similarly, Yang (2011) highlighted that music and songs are highly memorable and assist learners in retaining speech patterns and vocabulary. According to Nation (1990), repetition supports vocabulary learning by enabling learners to recall lexical items more easily. Because many required vocabulary items consist of multiword units rather than single words, songs can serve as efficient tools for helping children learn lexical patterns that can later be recalled naturally during spoken interaction (Murphey, 1992b).

A number of studies (e.g., Boustani & Al Abdwani, 2023; Chang, 2024) have explored the effects of music on memory, examining whether songs contribute to improved recall of newly learned vocabulary items. These studies have shown that music can enhance L2 learning when information is presented alongside musical input. Research has further demonstrated the benefits of music for memory performance (Ferreri & Verga, 2016; Musliu et al., 2017). Medina (2002) also found that when songs were used for vocabulary learning, students were more relaxed, receptive to new words, and able to recall them more easily.

Extensive research has confirmed the positive role of songs in L2 vocabulary development. Li and Brand (2009) compared full, partial, and no exposure to songs among Chinese ESL learners and found that learners receiving full exposure performed best, whereas the partial-exposure group performed worst, suggesting that inconsistent musical input may hinder learning. Similarly, Alipour et al. (2012) examined 60 Iranian EFL learners randomly assigned to musical and nonmusical instruction. Using posttests based on rap, pop, and country songs, they found that the music group significantly outperformed the nonmusic group in vocabulary retention and recall, demonstrating that songs are both engaging and pedagogically beneficial.

Complementing these findings, Schwarz (2013) reported that Austrian learners' incidental exposure to pop songs outside school facilitated implicit vocabulary learning. More recently, Zhang et al. (2023) found that adolescent Chinese ESL learners who learned song lyrics through singing outperformed those who learned through recitation in vocabulary recall and pronunciation, supporting the value of singing familiar melodies in L2 instruction.

Storytelling has also been widely recognized as an effective strategy for vocabulary acquisition in EFL instruction. Theoretically, stories provide meaningful contexts for introducing new words, promote prediction and inference, and serve as scaffolding for constructing and reconstructing linguistic knowledge (Wasik & Bond, 2001; Wright, 2000; Gibbons, 2002, as cited in Ioannou-Georgiou & Ramírez-Verdugo, 2010). Ellis and Brewster (1991) further emphasized its pedagogical value, noting that storytelling motivates learners, enhances social interaction, strengthens listening and concentration skills, and allows teachers to adapt vocabulary and sentence structures flexibly. Empirical evidence across age groups supports these claims. Abasi and Soori (2013) demonstrated that storytelling significantly improved vocabulary acquisition among 20 Iranian kindergarteners through pre- and posttests using pictorial items from the instructional storybook. Extending this work, Nemat Tabrizi and Abbasi (2016) found that generating oral summaries of short stories over 20 sessions significantly enhanced vocabulary learning among 75 Iranian intermediate students, with no gender differences observed. Similarly, Amin Afshar and Mojavezi (2017) reported that, among 50 intermediate female students aged 15–16, visual storytelling led to significantly better vocabulary acquisition and retention than aural storytelling after three months of instruction and a delayed posttest. More recently, Nami and Asadnia (2024) demonstrated that collaborative learner-generated digital stories enhanced vocabulary learning and cognitive-affective engagement among undergraduate EFL learners through multimodal and contextualized tasks. Likewise, Chandy et al. (2025) found that repeated exposure to narrative

text increased learners' processing speed and short-term vocabulary acquisition, underscoring the importance of effective story-based instructional design.

In a related study, Albaladejo et al. (2018) examined the effects of songs, stories, and combined song-story input on vocabulary development among seventeen children aged two to three. The children were exposed to 15 target lexical items through each condition, and behavioral learning patterns were analyzed via video recordings. Results showed that the story-only group achieved the highest vocabulary gains, while the song-only group performed least successfully, highlighting implications for incidental vocabulary acquisition in preschool EFL settings.

Research comparing songs and stories remains limited. Although numerous studies have investigated these methods among monolingual and ESL learners (Collins, 2010; Joyce, 2011; Lugo-Neris et al., 2010; Medina, 1990), fewer studies have focused specifically on EFL learners. In addition, variations in class size complicate generalizations, with some interventions conducted in small classrooms (Collins, 2010; Leśniewska & Pichette, 2014; Lin, 2014) and others in much larger groups (Coyle & Gómez Gracia, 2014).

Gender differences may also influence vocabulary learning and retention. Sylwester (1995) noted that the corpus callosum, which supports interhemispheric communication, tends to be larger in females, potentially enabling faster processing. Accordingly, a more developed corpus callosum may facilitate quicker integration of lexical processing in the left hemisphere with conceptual understanding in the right hemisphere (Uster, 2008). Supporting this claim, El-Nahhal (2011) found that female fourth-grade L2 learners outperformed males in vocabulary learning when songs were used as instructional material.

Overall, although substantial research (e.g., Akbarpour, 2014; Yilmaz, 2017) has examined gender differences in the use of strategies and techniques for L2 vocabulary acquisition, there remains a scarcity of studies investigating how male and female EFL learners develop

vocabulary through song- and story-based instruction. Therefore, the present study aimed to explore whether significant differences exist between male and female learners in vocabulary development when exposed to song- and story-based instructional methods.

Methodology

Participants

The study involved 60 elementary-level EFL learners, both male and female, aged 8 to 12, enrolled in two language institutes in Isfahan province, Iran. All participants were native Persian speakers and were divided into four groups: two song-based groups and two story-based groups, each comprising 15 male and 15 female learners. Participants' proficiency was confirmed as elementary through the Oxford Quick Placement Test (OQPT). During the intervention, learners acquired basic vocabulary related to routines, animals, and similar topics through songs or stories. Classes lasted 30 minutes, twice a week. This study aimed to examine which instructional approach was more effective for vocabulary acquisition and retention among elementary-level male and female EFL learners.

Instruments

To determine participants' proficiency and ensure group homogeneity, the OQPT was administered at the outset of the study. The OQPT has established reliability ($r = .90$) and validity and consists of 60 multiple-choice items assessing grammar, vocabulary, and reading comprehension.

To measure vocabulary learning and retention, three researcher-developed vocabulary assessments—a pretest, a posttest, and a delayed posttest—were employed for both the song and story groups. Each assessment contained 20 target lexical items drawn from the instructional materials. The pretest was administered prior to the treatment to confirm that the participants lacked prior knowledge of the lexical items included in the instructional sessions.

This ensured that any subsequent gains could be attributed to the intervention rather than familiarity with the vocabulary. The posttest was administered immediately after the completion of the instructional sessions to measure learners' immediate vocabulary gains. The delayed posttest, administered two weeks later, assessed the participants' retention of the target lexical items over time. Although the three assessments contained the same target words, the posttest and delayed posttest were reshuffled versions of the pretest to reduce the likelihood of recall based on item order rather than actual learning.

Each test included a range of item formats—multiple-choice questions, true/false items, fill-in-the-blank exercises, and matching tasks—to provide a comprehensive evaluation of both receptive and productive lexical knowledge. The internal consistency of the vocabulary assessments was examined using Cronbach's alpha and yielded a reliability estimate of .74, which is considered acceptable for research purposes. To establish content validity, the tests were reviewed by three experienced EFL instructors. Their recommendations regarding wording, item clarity, and alignment with learners' proficiency levels were incorporated into the final versions.

Regarding the instructional materials, five songs and five stories were selected from *Family and Friends 2* (Simmons, 2014). The selected songs were "If You're Tired and You Know It," "Places to Go," "We're Getting Ready for the Wedding," "On the Farm," and "The House Was So Untidy." The selected stories were "They're Happy Now," "Where Does She Work?", "What Are You Wearing?", "Look at All the Animals," and "Look at the Photos." All materials were presented using accompanying audio tracks and visual supports projected through data-show hardware.

Procedure

The study was conducted over a seven-week period with 60 elementary-level EFL learners aged 8 to 12. Participants were recruited from four intact classes in the same language institute and assigned to four groups according to gender and instructional condition: two song-based groups, each with 15 male and 15 female learners, and two story-based groups, each with 15 male and 15 female learners. Because intact classes were used, a quasi-experimental design was employed without randomization. Group homogeneity was verified using the OQPT, and only learners scoring within the elementary-level range were included. This procedure ensured that differences in vocabulary learning and retention could be attributed to the instructional interventions rather than pre-existing proficiency differences.

Prior to instruction, all participants completed a researcher-developed vocabulary pretest to establish baseline knowledge and confirm unfamiliarity with the target lexical items. The pretest included multiple formats, such as multiple-choice, true/false, fill-in-the-blank, and matching exercises, enabling assessment of both receptive and productive knowledge. Establishing baseline knowledge ensured that any subsequent gains could be attributed to incidental exposure through the instructional activities rather than prior familiarity. The instructional phase consisted of five sessions for each group, during which 50 lexical items were introduced incidentally, five per session, without explicit verbal explanation. Each session lasted approximately 40 minutes and incorporated multi-modal input, including audio, visual, and interactive components, to maximize engagement and comprehension.

In the song-based sessions, vocabulary was introduced incidentally through meaningful exposure to the songs. Learners first observed illustrations accompanying the song and responded to guided questions designed to activate prior knowledge and encourage predictions about the content. The song was then played twice, allowing learners to familiarize themselves with its melody, rhythm, and lyrics. Learners repeated the song phrase by phrase, noticing the

meaning of new words through the accompanying visuals. This was followed by singing the entire song aloud and reading the lyrics collectively and individually. These activities reinforced vocabulary recognition, pronunciation, rhythm, and fluency. Learners were encouraged to engage with the songs outside class and complete related exercises in the *Family and Friends 2* textbook (Simmons, 2014), providing repeated exposure that strengthened incidental acquisition and retention. By integrating music, visuals, and interactive practice, the song-based sessions created a rich environment for incidental vocabulary learning.

The story-based sessions also emphasized incidental learning, focusing on comprehension and contextualized understanding. Learners first examined the illustrations accompanying the story and provided brief oral comments to activate prior knowledge. The animated story was presented, and learners repeated it phrase by phrase, noticing word meanings in context. The teacher then read the printed story aloud while directing attention to the visuals, allowing learners to infer unfamiliar vocabulary naturally. To consolidate learning and encourage active production, learners memorized key dialogues and performed role-plays in pairs or small groups. This combination of visual, auditory, and interactive elements exposed learners to the lexical items in varied, meaningful contexts, supporting both incidental acquisition and communicative competence.

Following the instructional sessions, participants completed a posttest to measure immediate vocabulary acquisition. A delayed posttest was administered two weeks later to assess retention. Throughout both song- and story-based sessions, the teacher maintained consistent procedures and provided guidance only to support comprehension and participation, avoiding explicit instruction of the target vocabulary. This structured, multi-modal approach ensured that vocabulary learning occurred incidentally. The pretest, posttest, and delayed posttest results accurately reflected the effects of songs and stories on vocabulary acquisition and retention, as well as the moderating role of gender.

Data Analysis

The Statistical Package for Social Sciences (SPSS, version 21) was applied to scrutinize the data. For research question # 1, one-way repeated measures ANOVA was employed to compare the song group's performance on the pretest, posttest, and delayed posttest. The same statistical procedure was used for research question # 2 to track the changes the story group went through over time on the pretest, posttest, and delayed posttest. For research question # 3, because there were two independent variables (i.e., gender and song/story), the aim was to investigate the effects of these two independent variables and the one dependent variable (i.e., vocabulary) and two-way ANCOVA was used.

Results

Song Group Participants' Vocabulary Knowledge

To reach the first goal of the study and to answer research question # 1, the vocabulary pretest, posttest, and delayed posttest scores of the song group were compared via a one-way repeated measures ANOVA to check the changes this group went through in the course of the experimentation (see Table 1 for the results of the descriptive statistics for this analysis):

Table 1

Descriptive Statistics for Pretest, Posttest, and Delayed Posttest Scores of Song Group

Tests	Mean	Std. Deviation	N	Skewness	Kurtosis
Pretest	2.93	1.83	30	-.14	-.57
Posttest	16.96	1.27	30	.11	.07
Delayed Posttest	16.36	.96	30	.02	-.04

As indicated in Table 1, the song group participants' mean score for the pretest ($M = 2.93$) was smaller than their posttest mean score ($M = 16.96$), which was slightly higher than their delayed posttest mean score ($M = 16.36$). By the way, the skewness and kurtosis values for these distributions indicate that they were neither skewed nor peaked, but normally distributed because the values were found to be less than ± 2.00 . To discover whether or not the differences

among the three mean scores of the song group participants were statistically substantial, the p value under the Sig. column in Table 2 should be checked:

Table 2

One-Way Repeated Measures ANOVA for Pretest, Posttest, and Delayed Posttest Scores of Song Group

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Pillai's Trace	.98	744.26	2.00	28.00	.00	.98
Wilks' Lambda	.01	744.26	2.00	28.00	.00	.98
Hotelling's Trace	53.16	744.26	2.00	28.00	.00	.98
Roy's Largest Root	53.16	744.26	2.00	28.00	.00	.98

As seen in Table 2, the p value for the Wilks' Lambda test was less than the significance level (i.e., $.00 < .05$), meaning that there was a statistically substantial difference between the song group participants' mean scores on the vocabulary pretest, posttest, and delayed posttest. To locate where exactly the differences laid, Table 3 should be checked:

Table 3

Post-Hoc Test for Pretest, Posttest, and Delayed Posttest Scores of Song Group

Time		Mean Difference	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Pretest	Posttest	-14.03*	.41	.00	-15.09	-12.97
	Delayed Posttest	-13.43*	.34	.00	-14.30	-12.55
Posttest	Pretest	14.03*	.41	.00	12.97	15.09
	Delayed Posttest	.60	.25	.07	-.05	1.25
Delayed Posttest	Pretest	13.43*	.34	.00	12.55	14.30
	Posttest	-.60	.25	.07	-1.25	.05

Table 3 illustrates a statistically significant improvement in the song group's vocabulary performance from the pretest ($M = 2.93$) to the posttest ($M = 16.96$), as indicated by a p value of .00. Similarly, the comparison between the pretest and the delayed posttest ($M = 16.36$) also yielded a statistically meaningful result, given that the p value was below the threshold of significance ($.00 < .05$). In contrast, the observed difference between the posttest and delayed posttest scores was not statistically significant ($p = .07$), as it exceeded the .05 level. These findings suggest that integrating songs into vocabulary instruction substantially enhanced

learners' vocabulary acquisition. Furthermore, the consistently higher posttest and delayed posttest mean scores, relative to the pretest, imply that the instructional effects were sustained over time, despite the minor reduction observed between the two later assessments. In sum, teaching vocabulary through songs markedly improved EFL learners' vocabulary knowledge and demonstrated a degree of durability over time.

Story Group Participants' Vocabulary Knowledge

To examine the progression of the story group participants' vocabulary performance throughout the study, a one-way repeated measures ANOVA was conducted, comparing their scores on the vocabulary pretest, posttest, and delayed posttest. Table 4 presents the descriptive statistics derived from the one-way ANOVA analysis:

Table 4

Descriptive Statistics for Pretest, Posttest, and Delayed Posttest Scores of Story Group

Tests	Mean	Std. Deviation	N	Skewness	Kurtosis
Pretest	3.23	1.61	30	.17	-.98
Posttest	15.73	1.41	30	-.25	-.09
Delayed Posttest	14.06	1.11	30	-.14	-.57

The mean score of the story group on the pretest ($M = 3.23$) was lower than both their posttest ($M = 15.73$) and delayed posttest ($M = 14.06$) mean scores. To determine whether these differences were statistically significant, it was necessary to consult the p value listed in the Sig. column of the one-way repeated measures ANOVA table (Table 5). However, prior to analyzing the ANOVA results, the assumption of normality needed to be verified. This was done by examining the skewness and kurtosis values reported in Table 4. Since all values fell within the acceptable range of ± 2 , the distributions of the story group's pretest, posttest, and delayed posttest scores were deemed normal. With this assumption satisfied, the analysis of the ANOVA results could proceed:

Table 5

One-Way Repeated Measures ANOVA for Pretest, Posttest, and Delayed Posttest Scores of Story Group

	Value	<i>F</i>	Hypothesis <i>df</i>	Error <i>df</i>	Sig.	Partial Eta Squared
Pillai's Trace	.98	736.71	2.00	28.00	.00	.98
Wilks' Lambda	.01	736.71	2.00	28.00	.00	.98
Hotelling's Trace	52.62	736.71	2.00	28.00	.00	.98
Roy's Largest Root	52.62	736.71	2.00	28.00	.00	.98

Casting a look at the *p* value in front of Wilks' Lambda (which is the most widely reported test in one-way repeated measures ANOVA) under the Sig. column reveals that the *p* value was below the threshold for significance ($.00 < .05$). This outcome signifies that there was a statistically significant difference among the story group's vocabulary scores across the pretest, posttest, and delayed posttest. To identify where these differences specifically occurred—whether between the pretest and posttest, the pretest and delayed posttest, or the posttest and delayed posttest—it is necessary to check the results presented in the post-hoc test table (see Table 6):

Table 6

Post-Hoc Test for Pretest, Posttest, and Delayed Posttest Scores of Story Group

Time		Mean Difference	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Pretest	Posttest	-12.50*	.35	.00	-13.39	-11.60
	Delayed Posttest	-10.83*	.32	.00	-11.65	-10.01
Posttest	Pretest	12.50*	.35	.00	11.60	13.39
	Delayed Posttest	1.66*	.31	.00	.87	2.45
Delayed Posttest	Pretest	10.83*	.32	.00	10.01	11.65
	Posttest	-1.66*	.31	.00	-2.45	-.87

The difference between the pretest ($M = 3.23$) and posttest ($M = 15.73$) mean scores of the story group participants did reach statistical significance, as the *p* value corresponding to this comparison equaled .00 that is less than .05. Moreover, the difference between the posttest and delayed posttest ($M = 14.06$) mean scores was found to be of statistical significance, as the related *p* value was found to be less than the significance level (i.e., $.00 < .05$). Finally, the difference between the pretest and delayed posttest scores was statistically significant ($p < .05$).

Hence, it could be construed that using stories could significantly improve the EFL participants' vocabulary knowledge, although this gain in knowledge may fall dramatically after a lapse of time. Table 6 also shows that the posttest mean score of the story group participants was substantially higher than their pretest mean score. Although their delayed posttest mean score dropped considerably, it was still far larger than the pretest mean score. This result, thus, confirms the contributory short-term and long-term effects of using stories for the purpose of improving EFL learners' vocabulary knowledge. Now that it was found that both songs and stories were effective tools of boosting EFL learners' vocabulary knowledge, it is high time we compared these two teaching techniques and examine the moderating role of gender in this regard.

Songs vs. Stories: The Moderating Role of Gender

Vocabulary Learning

In order to see whether stories or songs had more impacts on the (fe)male EFL participants' vocabulary knowledge, a comparison of the participants' posttest scores across the two instructional groups was necessary. For this purpose, a two-way ANCOVA was conducted, the results of which are presented in Tables 7 and 8:

Table 7

Descriptive Statistics for Posttest Scores of (Fe)male Participants in Two Groups

Groups	Gender	Mean	Std. Deviation	N
Story	Male	15.74	1.57	15
	Female	15.73	1.27	15
	Total	15.73	1.41	30
Song	Male	17.00	1.25	15
	Female	16.93	1.33	15
	Total	16.96	1.27	30
Total	Male	16.36	1.54	30
	Female	16.33	1.42	30
	Total	16.35	1.47	60

The mean vocabulary posttest scores for male and female participants in the story group were 15.74 and 15.73, respectively. In comparison, male and female learners in the song group achieved mean scores of 17.00 and 16.93, respectively. Moreover, the overall mean score for the story group was 15.73, whereas the song group had a higher total mean of 16.96. To assess whether these observed differences between genders and instructional groups were statistically significant, the p values corresponding to Groups and Gender in the Sig. column of the two-way ANCOVA output needed to be reviewed (see Table 8):

Table 8

Two-Way ANCOVA for Posttest Scores of (Fe)male Participants in Two Groups

Source	Type III Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	Sig.	Partial Eta Squared
Corrected Model	23.37	4	5.84	3.08	.02	.18
Intercept	3536.99	1	3536.99	1865.59	.00	.97
Pretest	.52	1	.52	.27	.60	.00
Groups	23.25	1	23.25	12.26	.001	.18
Gender	.004	1	.004	.002	.96	.00
Groups * Gender	7.68	1	7.68	.00	.99	.00
Error	104.27	55	1.89			
Total	16167.00	60				
Corrected Total	127.65	59				

As indicated in Table 8, a statistically significant difference was observed between the vocabulary posttest scores of participants in the story group ($M = 15.73$) and those in the song group ($M = 16.96$). because the p value under the Sig. column in front of Groups was smaller than the specified level of significance (i.e., $.001 < .05$). The magnitude of this difference, as shown under the Partial Eta Squared column, was very large based on Cohen (1988, as cited in Pallant, 2010): .01 = small, .06 = moderate, and .14 = large.

However, the p value associated with the Gender variable exceeded the significance threshold ($.96 > .05$), suggesting that gender did not influence the relationship between the type of instruction—songs versus stories—and vocabulary acquisition. Additionally, the interaction effect between the two independent variables, namely instructional method (songs/stories) and gender, was not statistically significant, as evidenced by the p value for Groups*Gender ($.99 >$

.05), which also surpassed the conventional significance level. As presented in Table 8, both male and female participants in the song group achieved notably higher vocabulary posttest scores compared to their counterparts in the story group, while the performance differences between males and females within each group were minimal.

Vocabulary Retention

To investigate the long-term impact of stories and songs on the vocabulary knowledge of male and female EFL learners, a second two-way ANCOVA was conducted to compare the participants' delayed posttest scores across the two groups. The findings of this analysis are presented in Tables 9 and 10:

Table 9

Descriptive Statistics for Delayed Posttest Scores of (Fe)male Participants in Two Groups

Groups	Gender	Mean	Std. Deviation	N
Story	Male	13.93	1.22	15
	Female	14.20	1.01	15
	Total	14.06	1.11	30
Song	Male	16.13	.99	15
	Female	16.60	.91	15
	Total	16.36	.96	30
Total	Male	15.03	1.56	30
	Female	15.40	1.54	30
	Total	15.21	1.55	60

In the story group, male and female participants achieved mean scores of 13.93 and 14.20, respectively, on the delayed posttest. In contrast, the mean scores for male and female learners in the song group were notably higher, at 16.13 and 16.60, respectively. These results indicate substantial differences between the two instructional groups—songs and stories—for both male and female participants. However, the overall mean scores for males ($M = 15.03$) and females ($M = 15.40$) across both groups did not reflect a considerable gender-based difference. To determine whether the differences between instructional methods and genders on the delayed posttest were statistically significant, it was necessary to examine the two-way ANCOVA output provided in Table 10:

Table 10

Two-Way ANCOVA for Delayed Posttest Scores of (Fe)male Participants in Two Groups

Source	Type III Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	Sig.	Partial Eta Squared
Corrected Model	84.95	4	21.23	20.40	.00	.59
Intercept	2949.87	1	2949.87	2834.80	.00	.98
Pretest	3.43	1	3.43	3.30	.07	.05
Groups	81.69	1	81.69	78.50	.00	.58
Gender	2.49	1	2.49	2.40	.12	.04
Groups * Gender	.52	1	.52	.50	.47	.009
Error	57.23	55	1.04			
Total	14035.00	60				
Corrected Total	142.18	59				

As presented in Table 10, a significant difference was found between the delayed posttest scores of the story group ($M = 14.06$) and the song group ($M = 16.36$), as the p value associated with Groups was below the alpha level ($.00 < .05$). The corresponding effect size, reported under the Partial Eta Squared column, was .58, indicating a very large effect for the type of instructional method—songs versus stories. In contrast, the p value linked to Gender ($.12 > .05$) exceeded the significance threshold, suggesting that gender did not exert a statistically significant influence on participants' delayed posttest performance. Additionally, the interaction between instructional method and gender was not statistically significant, as evidenced by a p value of .47 for Groups* Gender, which was also greater than the .05 level.

Also, Table 10 indicates that, in both story and song groups, the females were slightly better than the males; yet, the differences between them were not considerable. In other words, the males and females in the story group were not substantially different from each other on the delayed posttest of lexicon. Similarly, the (fe)male participants in the song group were roughly similar in terms of their delayed posttest scores. Nonetheless, the song group participants managed to significantly outperform the story group participants on the delayed posttest of vocabulary, giving rise to the long-term effectiveness and superiority of songs over stories when it comes to vocabulary retention.

Discussion

The present study investigated the effectiveness of songs and stories in developing the vocabulary knowledge of Iranian elementary-level EFL learners and examined whether gender played a moderating role in these outcomes. The results revealed several important insights that contribute to the growing body of research on multimodal vocabulary instruction and incidental lexical learning.

The findings demonstrated that songs were a highly effective medium for promoting vocabulary learning and retention. Learners who received song-based instruction consistently outperformed those who learned vocabulary through stories, showing both stronger immediate gains and more stable retention over time. This superiority is consistent with earlier studies highlighting the pedagogical benefits of songs for lexical development (Alipour et al., 2012; Medina, 1993, 2002; Murphey, 1992a, 1992b). Songs provide natural repetition, predictable rhythmic patterns, and melodic cues, all of which strengthen memory encoding and retrieval (Ferreri & Verga, 2016; Musliu et al., 2017). The affective benefits of music further support learning by lowering anxiety and increasing engagement (Yang, 2011; Schön et al., 2004). More recent research also confirms the advantage of musical input for L2 vocabulary learning, with studies showing improved recall, pronunciation, and retention when learners sing or listen to accompanied text (Zhang et al., 2023; Chang, 2024; Boustani & Al Abdwani, 2023). Collectively, these findings affirm that songs offer a powerful and engaging pathway for incidental vocabulary learning among young EFL learners.

Although stories were not as effective as songs in sustaining long-term retention, they still produced meaningful vocabulary gains. Story-based instruction provided learners with rich and meaningful contexts, allowing them to infer meanings from narrative events and situational cues. This aligns with theoretical claims that stories facilitate lexical development through scaffolding, prediction, and contextualization (Wasik & Bond, 2001; Wright, 2000; Gibbons,

2002, as cited in Ioannou-Georgiou & Ramírez-Verdugo, 2010). Empirical evidence from studies on Iranian learners likewise supports the positive influence of storytelling on vocabulary learning (Abasi & Soori, 2013; Nemat Tabrizi & Abbasi, 2016). Moreover, multimodal approaches to storytelling—such as visual or digital narratives—have been shown to enhance retention, suggesting that additional supports can strengthen vocabulary outcomes (Amin Afshar & Mojavezi, 2017; Nami & Asadnia, 2024). The decline in retention found in this study, however, echoes research indicating that story-based lexical gains may diminish without reinforcement (Chandy et al., 2025). Thus, while stories are effective for initial vocabulary development, their long-term impact may require repeated or multimodal exposure.

The direct comparison between the two instructional methods demonstrated a clear advantage for songs. Unlike stories, which often introduce vocabulary in varied narrative contexts, songs rely on structured repetition, predictable patterns, and memorable rhythms that strengthen phonological and semantic encoding. This may explain why learners in the song condition tended to retain vocabulary more consistently. These findings differ from those of Albaladejo et al. (2018), who reported stronger vocabulary gains in story-based conditions among preschool EFL learners. Differences in learners' ages, cognitive developmental stages, and the design of instructional materials may account for these discrepancies. The present results contribute to addressing a gap noted in the literature: although songs and stories have often been studied separately, few studies have directly compared their effects within an EFL context, particularly with elementary learners. The findings suggest that while both methods support vocabulary learning, songs may provide a more efficient mechanism for ensuring both acquisition and retention.

Another key finding was that gender did not exert a significant effect on vocabulary performance, nor did it interact with the instructional method. Male and female learners benefited similarly from both songs and stories. Although female learners showed a slight

retention advantage, the difference was not statistically meaningful. These results contrast with earlier claims that gender differences in brain structure or cognitive processing may influence vocabulary learning (Sylwester, 1995; Uster, 2008; El-Nahhal, 2011). However, they are in line with studies reporting no significant gender-based differences in vocabulary acquisition through storytelling or other communicative methods (Nemat Tabrizi & Abbasi, 2016; Yilmaz, 2017). The present findings suggest that at the elementary level, vocabulary learning through multimodal input may override potential gender-related differences, enabling equitable learning outcomes for all learners.

Conclusion

This study examined the effectiveness of songs and stories in promoting vocabulary learning and retention among Iranian elementary-level EFL learners, while also exploring whether gender moderated these outcomes. The findings confirmed that both instructional methods significantly enhanced learners' vocabulary knowledge, underscoring their pedagogical value in early language education. However, the comparative analysis revealed that songs were more effective than stories, particularly in sustaining long-term retention. Gender did not exert a significant influence, indicating that both male and female learners benefited equally from the interventions.

By directly comparing songs and stories, the study contributes to filling a gap in the literature, as most prior research examined these methods in isolation. The results highlight the importance of selecting instructional techniques that not only facilitate immediate vocabulary acquisition but also support durable retention. Pedagogically, the findings suggest that songs should be prioritized in elementary EFL classrooms, while stories remain valuable for contextualized learning and engagement. Curriculum designers and textbook developers are

encouraged to integrate both methods into instructional materials, with greater emphasis on songs to maximize vocabulary outcomes for young learners.

While the study offers meaningful insights, its limitations should be acknowledged. The relatively small sample size and focus on Iranian elementary learners may constrain generalizability. Future research could replicate the study across larger and more diverse populations, explore hybrid approaches that combine songs and stories, and investigate digital adaptations to reinforce vocabulary retention. Longitudinal studies would also be valuable in assessing how these methods influence vocabulary development over extended periods.

In sum, this research demonstrates that songs and stories are effective tools for enhancing vocabulary learning among elementary EFL learners, with songs offering a distinct advantage in retention. The absence of gender differences further confirms the broad applicability of these methods. By providing empirical evidence and practical recommendations, the study contributes to advancing vocabulary pedagogy in EFL contexts and opens avenues for future exploration into innovative instructional practices.

References

- Abasi, M., & Soori, A. (2014). Is storytelling effective in improving the English vocabulary learning among Iranian children in kindergartens? *International Journal of Education and Literacy Studies*, 2(3), 7-11.
- Abdollahzadeh, E., & Amiri-Vardani, N. (2010). The effect of vocabulary instruction through semantic mapping among EFL learners with different perceptual learning. *Journal of English Language Studies*, 1(2), 1-22.
- Akbarpour, S. (2014). *The effect of using songs on young Iranian EFL learners' vocabulary learning: A case of gender* [Unpublished master's thesis]. Shahrekord University.
- Albaladejo, S., Coyle, Y., & Roca de Larios, J. (2018). Songs, stories, and vocabulary acquisition in preschool learners of English as a foreign language. *System*, 76, 116-128.
- Alharthi, M. (2025). Gender-based differences in receptive vocabulary knowledge: A study of English achievement test performance among Saudi EFL undergraduates. *International Journal of English Language Education*, 13(1), 41-61.
- Alipour, M., Gorjian, B., & Zafari, I. (2012). The effects of songs on EFL learners' vocabulary recall and retention: The case of gender. *Advances in Digital Multimedia*, 1(3), 140-143.
- Amin Afshar, M., & Mojavezi, A. (2017). The effect of aural and visual storytelling on vocabulary retention of Iranian EFL learners. *English Language Teaching*, 10(4), 92-99.
- Ayotte, S. B. (2004). *The acquisition of verb forms through song* [Doctoral dissertation, McGill University]. *Dissertation Abstracts International*, 65(9), 3356-A.

- Boustani, N., & Al Abdwani, T. (2023). The effect of music on communication ability: A case of vocabularies. *Journal of Business, Communication & Technology*, 2(1), 1-12.
- Brown, H. D. (2007). *Teaching by principles: An interactive approach to language pedagogy* (3rd ed.). Pearson Education.
- Carter, R., & McCarthy, M. (2014). *Vocabulary and language teaching*. Routledge.
- Celce-Murcia, M. (Ed.). (2001). *Teaching English as a second or foreign language* (3rd ed.). Heinle & Heinle.
- Chandy, R., Serrano, R., & Pellicer-Sánchez, A. (2025). The effect of context on the processing and learning of novel L2 vocabulary while reading. *Applied Psycholinguistics*, 45(6), 1086–1113.
- Chang, A.C.S. (2024). The effects of varying practice modes on L2 learners' vocabulary retention through songs: Listening, singing, and oral reading. *The Electronic Journal for English as a Second Language*, 28(3), 1-18.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Collins, M. F. (2010). ELL preschoolers' English vocabulary acquisition from storybook reading. *Early Childhood Research Quarterly*, 25(1), 84-97.
- Coyle, Y., & Gómez Gracia, R. (2014). Using songs to enhance L2 vocabulary acquisition in preschool children. *ELT Journal*, 68(3), 276-285.
- Ellis, G., & Brewster, J. (1991). *The storytelling handbook for primary teachers of English*. Penguin Books.
- Ellis, G., & Brewster, J. (2002). *Tell it again! The storytelling handbook for primary English language teachers* (2nd ed.). British Council.
- El-Nahhal, M. M. (2011). *The effectiveness of using children songs on developing the fourth graders' English vocabulary in Rafah governmental schools* [Unpublished master's thesis]. Al-Azhar University.
- 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.
- Fischler, J. (2006). *The rap on stress: Instruction of word and sentence stress through rap music* [Unpublished master's thesis]. Hamline University.
- Gibbons, P. (2002). *Scaffolding language, scaffolding learning: Teaching second language learners in the mainstream classroom*. Heinemann.
- Heidari, K. (2019). Willingness to communicate: A predictor of pushing vocabulary knowledge from receptive to productive. *Journal of Psycholinguistic Research*, 48, 903-920.
- Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *The Canadian Modern Language Review*, 61(3), 355-382.
- Ioannou-Georgiou, S., & Ramírez-Verdugo, M. D. (2010). *Stories as a tool for teaching and learning in CLIL*. Cyprus Pedagogical Institute.
- Joyce, M. (2011). *Vocabulary acquisition with kindergarten children using song picture books* [Unpublished doctoral dissertation]. Northeastern University.
- Krashen, S. D. (1981). *Second language acquisition and second language learning*. Pergamon Press.
- Leśniewska, J., & Pichette, F. (2014). Songs vs. stories: Impact of input sources on ESL vocabulary acquisition by preliterate children. *International Journal of Bilingual Education and Bilingualism*, 19(1), 18-34.
- Li, X., & Brand, M. (2009). Effectiveness of music on vocabulary acquisition, language usage, and meaning for Mainland Chinese ESL learners. *Contributions to Music Education* 36, 73-84.
- Lin, L. C. (2014). Learning word meanings from teachers' repeated story read-aloud in EFL primary classrooms. *English Language Teaching*, 7(7), 68-81.

- Lo, R., & Li, H. (1998). Songs enhance learner involvement. *English Teaching Forum*, 36(3), 8-11.
- Lugo-Neris, M., Wood Jackson, C., & Goldstein, H. (2010). Facilitating vocabulary acquisition of young English language learners. *Language, Speech, and Hearing Services in Schools*, 41(3), 314-327.
- McLaren, P., & Lankshear, C. (1994). *Politics of liberation: Paths from Freire*. Routledge.
- Medina, S. L. (1993). The effect of music on second language vocabulary acquisition. *FEIS News: National Network for Early Language Learning*, 6(3), 1-8.
- Medina, S. L. (2002). Using music to enhance second language acquisition: From theory to practice. In J. Lalas & S. Lee (Eds.), *Language, literacy, and academic development for English language learners* (pp. 115-126). Pearson.
- Murphey, T. (1992a). *Music and song*. Oxford University Press.
- Murphey, T. (1992b). The discourse of pop songs. *TESOL Quarterly*, 26(4), 770-774.
- Musliu, A., Berisha, B., Musaj, A., Latifi, D., & Peci, D. (2017). The impact of music on memory. *European Journal of Social Sciences Education and Research*, 4(4), 222-227.
- Nami, F., & Asadnia, F. (2024). Exploring the effect of EFL students' self-made digital stories on their vocabulary learning. *System*, 120, 103205.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Newbury House.
- Nation, I. S. P. (2022). *Learning vocabulary in another language* (3rd edition). Cambridge University Press.
- Nemat Tabrizi, A. R., & Abbasi, S. (2016). The effect of oral summary of short stories on Iranian intermediate EFL learners' vocabulary learning: With a focus on gender. *International Journal of English Linguistics*, 6(6), 111-120.
- Neumann, S. B. (2004). Hear, hear. *Scholastic Parent & Child*, 11(4), 22-23.
- Nyikos, M., & Fan, M. (2007). A review of vocabulary learning strategies: Focus on language proficiency and learner voice. In A. D. Cohen & E. Macaro (Eds.), *Language learner strategies* (pp. 251-274). Oxford University Press.
- Pallant, J. (2010). *SPSS survival manual: A step by step guide to data analysis using SPSS* (4th ed.). Allen & Unwin.
- Penno, J. F., Wilkinson, I. A. G., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology*, 94(1), 23-33.
- Pourshahian, B., Yousefi, A. A., & Rezvani, K. A. (2012). Does applying vocabulary learning strategies vary based on gender? The case of Turkish EFL learners. *Research on Humanities and Social Sciences*, 2(4), 1-11.
- Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, 52(3), 513-536.
- Qian, D. D., & Lin, L. H. F. (2020). The relationship between vocabulary knowledge and language proficiency. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 66-80). Taylor & Francis Group.
- Ransdell, S. E., & Gilroy, L. (2001). The effects of background music on word processing writing. *Computers in Human Behavior*, 17(2), 141-148.
- Read, J. (2000). *Assessing vocabulary*. Cambridge University Press.
- Rivers, W. M., & Nunan, D. (1991). *Language teaching methodology: A textbook for teachers*. Prentice Hall International.
- Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329-363.
- Schoepp, K. (2001). Reasons for using songs in the ESL/EFL classroom. *The Internet TESL Journal*, 7(2). Retrieved from: <http://iteslj.org/Articles/Schoepp-Songs.html>

- Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2007). Songs as an aid for language acquisition. *Cognition*, 106(2), 975-983.
- Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, 41, 341-349.
- Schwarz, M. (2013). Learning with Lady GaGa & Co: Incidental EFL vocabulary acquisition from pop songs. *VIEWS*, 22, 17-48.
- Simmons, N. (2014). *Family and friends 2* (2nd ed.). Oxford University Press.
- Soleimani, H., & Akbari, M. (2013). The effect of storytelling on children's learning English vocabulary: A case in Iran. *International Research Journal of Applied and Basic Sciences*, 5(1), 104-113.
- Sylwester, R. (1995). *A celebration of neurons: An educator's guide to the human brain*. Association for Supervision and Curriculum Development.
- Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science*, 11(3), 188-195.
- Uster, S. (2008). *The role of brain-based gender differences on the vocabulary learning and consolidation skills and strategies* [Unpublished master's thesis]. Middle East Technical University.
- van Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: A dimension approach. *System*, 41(3), 609-624.
- Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, 61(1), 219-258.
- Wasik, B. A., & Bond, M. A. (2001). Beyond the pages of a book: Interactive book reading and language development in preschool classrooms. *Journal of Educational Psychology*, 93(2), 243-250.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46-65.
- Webb, S., Uchihara, T., & Yanagisawa, A. (2023). How effective is second language incidental vocabulary learning? A meta-analysis. *Language Teaching*, 56(2), 1-20.
- Wright, A. (2000). Stories and their importance in language teaching. *Humanising Language Teaching*, 2(5), 1-6.
- Yang, L. (2011). *Using music in English as a second language classroom* [Unpublished master's thesis]. University of Wisconsin-Platteville.
- Yilmaz, V. G. (2017). The role of gender and discipline in vocabulary learning strategy use of Turkish graduate EFL learners. *International Journal of Innovation and Research in Educational Sciences*, 4(1), 57-64.
- Zhang, X., Baills, F., & Prieto, P. (2023). Singing songs facilitates L2 pronunciation and vocabulary learning: A study with Chinese adolescent ESL learners. *Languages*, 8(3), 1-24.