8-17-2023 10:30 AM

The Effects of Song Use on Vocabulary Learning: Studies of Distribution of Practice, Modes of Input, Retrieval, and the Input-Output-Input Sequence of Exposure

Niousha Pavia, *Western University*

Supervisor: Faez, Farahnaz, *The University of Western Ontario*
Joint Supervisor: Boers, Frank, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Doctor of Education degree in Education

Abstract

This dissertation investigates the effects of the distribution of practice, modes of input, retrieval, and the input-output-input sequence of exposure on incidental vocabulary learning from songs. This thesis takes an integrated article format, organized into five chapters, including an introduction, articles one, two, and three, and a conclusion. The participants (*N*=225) across all three studies involved Thai students learning English as a foreign language (EFL) in Thailand. All three studies measured vocabulary learning gains by comparing participants' scores on the vocabulary knowledge tests used for pretests, immediate posttests, and two weeks delayed posttests. Study one partially replicated and extended Pavia et al.'s (2019) study investigating the effects of repeated listening to songs on incidental learning of single words and collocations with the addition of spaced listening conditions exploring the effects of distributed practice on vocabulary learning. The results indicated that repeated listening to songs in mass listening and spaced listening conditions may foster learning of single words and collocations. Furthermore, the mass listening condition produced higher learning gains on the immediate posttest. However, participants in the spaced practice condition showed less regression than those in the mass practice condition on the delayed posttest. Study two explored the effects of modes of input (i.e., Listening only (L), listening while reading the lyrics (LL), listening and singing (LS), and listening and singing while reading the lyrics (LSL)) on learning formulaic sequences (FS). Based on the results, it was found that the LSL group had the most effective mode of learning through songs. Study three examined the effect of retrieval and the sequence of input and output on incidental learning of receptive and productive aspects of FS from a song. The results indicated that using retrieval activity in the input-input-output-output-input sequence of exposure was the most effective condition for learning FS incidentally from a song. Overall,

the three studies in this volume provided empirical evidence that songs can be an invaluable source of language input for vocabulary acquisition and how learning can be optimized by manipulating the distribution of practice, modes of input, and retrieval activities.

Keywords: Incidental vocabulary learning; songs; replication and extension; massed listening; spaced listening; modes of input; retrieval, sequential practice; formulaic sequences, collocations, receptive knowledge, productive knowledge

Summary for Lay Audience

Students learning a second or foreign language have difficulty learning new words. Research in the past has established helpful tips for language learners to utilize various sources of language input. (i.e., Books, movies, songs, etc.). However, many factors have yet to be explored when using songs for language learning. Through three studies, this thesis investigated the impact of various techniques on language learning with songs. Language learning is measured by the number and aspects of words learned due to the interventions carried out through each study. Pavia et al. (2019) showed that listening to a song repeatedly can increase the likelihood of learning new words. However, the effect of the time interval between learning conditions was not clear. Thus, the first study in this thesis examined how learning conditions' timing and schedule affected learning outcomes. The results showed that listening in one session had higher learning gains in the short term. However, if the listening sessions were separated by time, more of what the students had learned was retained in the long term. The second study examined how combining listening with reading the song lyrics and singalong activities affected learning outcomes. The results showed that having the lyrics present and pushing learners to produce the

song through singing resulted in higher learning outcomes. The final study explored the effects of sequential exposure to a song by alternating listening and singing activities that included singing and writing the lyrics from memory. The learning conditions in this study ended with a final listening session allowing learners to consolidate their knowledge of the target language. The results demonstrated that the added recall activity allowed them to notice the gaps in their knowledge, so when they could listen again, they paid closer attention to the language. This resulted in higher learning compared to the other conditions. Overall, the three studies in this thesis provided further evidence that songs can foster language learning. The learning experience can be enhanced by taking advantage of changes in the practice schedule, using complementary resources such as song lyrics and activities such as recalling the song from memory.

Co-Authorship statement

The three articles presented in chapters two, three, and four will be submitted for publication in peer-reviewed journals as co-authored papers. For all three studies, I have been responsible for 100% of designing the procedure, creating data collection materials, recruiting participants, performing the study, collecting, and analyzing the data, and preparing the final manuscript. The contribution of the other authors included consultation and feedback throughout the project, from guidance on forming research questions to evaluating the validity of the instruments and interpretation of the results. The feedback was given in several forms, including in-manuscript commentary, in-person discussions, general comments on ideas, additional references for consideration, and editing suggestions. Nevertheless, I was solely responsible for reflecting on the feedback and deciding to incorporate them into the finalized manuscript.

Acknowledgements

I would like to extend my sincere gratitude to my supervisors, Dr. Farahnaz Faez and Dr. Frank Boers, for their unwavering support at every stage of this project. A special thank you also goes to my mentors, Dr.Stuart Webb and Dr. Paul Trembley, and my colleagues Dr. Su Kyung Kim, Dr. Takumi Uchihara, Dr. Akifumi Yanagisawa, Juliane Martini, and Emi Iwaizumi, for inspiring and supporting me throughout my graduate career at Western University. I would also like to thank the National Research Council of Thailand (NRCT) for allowing me to pursue this project in their extraordinary country and the Thai school, teachers, and students for their participation. Finally, I am forever grateful to my family for their support and encouragement, without which this project would not have been possible.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

MEG	Magnetoencephalography
L1	First language
L2	Second language
SLA	Second Language Acquisition
NRCT	National Research Council of Thailand
BNC	British National Corpus
COCA	Corpus of Contemporary American English
ESL	English a second language
EFL	English as a foreign language
MWU	Multi-word unites
FL	Formulaic Language
FS	Formulaic Sequences
TOPRA	Type of Processing-Resource Allocation
TFA	Technique Feature Analysis
C	Control Group
C1	Control Group 1
C2	Control Group 2
E1	Experimental Group 1
E2	Experimental Group 2
L	Listen only Group
LL	Listen while reading the lyrics Group
LS	Listen and sing Group

LSL	Listen and sing while reading the lyrics Group

Chapter 1: Introduction

This chapter provides the theoretical background and the motivation behind the three studies in this dissertation. A brief description of the context in which the data were collected, the ethical considerations, and the shared procedure between the three studies is also provided. Following this chapter, chapter two focuses on the first study, "Incidental Vocabulary Learning through Massed Versus Spaced Listening to Songs: Replication and Extension." Chapter three focuses on the second study, "Receptive and Productive Learning of Formulaic Sequences from Spaced Listening to Songs: The Role of Different Modes of Input." Chapter Four focuses on the third and final study of this dissertation, "The Effects of Retrieval on Incidental Vocabulary Learning from Songs". Chapter Five concludes this thesis with a summary of the finding, potential implications, and future direction for research.

1.1 Theoretical Background

English as lingua franca or the common language of communication between speakers of different languages dates to the 1960s, during which the English language began to expand worldwide (Richards, 2008). Consequently, non-English-speaking countries introduced English as a foreign language into their statutory school curriculum (Coyle & Gomez Gracia, 2014). This sparked more interest in second/foreign language learning research.

Much of the focus of research in language learning has been on learning from reading. (Nagy et al.,1985; Horst et al., 1998; Nation, 2015; Webb & Chang, 2015a, 2015b; Pellicer-Sánchez, 2017). In recent years research has also explored learning from listening to different spoken texts, including academic lectures (Vidal, 2003; 2011) and short passages (Van Zeeland & Schmitt, 2013). However, with the technological advances and access to the internet, creating different forms of entertainment and media use, fewer people read books and are less likely to be

exposed to the English language through reading every day. Reports from the European Commission (2002) investigating media use in 15 European Union member countries indicated that 40% of the respondents claimed not to read books at all. In North America, individuals tend to watch television (Statistics Canada, 1998; United States Department of Labor, 2006) and listen to music (Levey et al., 2011) more frequently than read books. Thus, it is not surprising that researchers are starting to advocate for further investigation of learning gains from sources of language input more frequently used by language learners, such as listening to songs and the use of musical compositions of language for second and foreign language acquisition (Pavia et al., 2019).

"Language and music define us as humans" (Patel, 2008, p.3). "Music is one of the most natural phenomena that inspire the brain and nervous system to feel, create, and move" (Randolph, 2017, p.35). Archeological excavations have found that musical instruments are among the oldest human-made artifacts (Levitin, 2007). In the most primitive human societies, in the absence of complex linguistic and cultural characteristics, where they do not have any fixed terms for colours or have a clear understanding of counting, they still create music in songs (Patel, 2008).

Language and music both involve sophisticated sound sequences. Every human being enters the world with two sound systems. The first is the linguistic sound system that allows for recognizing and producing vowels, consonants, and pitch contrasts associated with the native language. The second allows for the recognition and production of musical timbre and pitches of native music. Children can easily recognize their culture's music and specific linguistic sounds without musical or linguistic instructions. Linguistic and musical proficiency develop simultaneously. Thus, as individuals become proficient with one language, they also become

proficient with the sound sequences of that culture's music. As adults with high linguistic and musical proficiency in their native tongue, individuals tend to have difficulty recognizing the sounds in other languages. They also tend to find the sounds of foreign music out of pitch and less enjoyable. This phenomenon has led cognitive science and neuroscience researchers to explore the hidden connection between language and music. (Patel, 2008).

Brandt et al. (2012) described language as a special type of music (p. 1). They suggested that music can provide the foundation for language learning. The initial aspects of language learning by newborns are the musical features of that language, including stress patterns and prosody. Neuroimaging and behavioural studies suggest that the brain uses the same structure to process the rules and the processes that govern the structure of sentences in a language and musical syntax (Levitin & Menon, 2003; Patel, 2008). In two studies, the first investigating how children and adults extract words from connected speech and the second exploring how children and adults extract words from songs, researchers found that similar learning mechanisms operate in both situations (Saffran et al., 1999; Pelucchi et al., 2009).

In addition, in a study using magnetoencephalography (MEG), researchers could localize the analysis of harmonic sequencing in the brain. The results indicated that Broca's area and its right-hemisphere homologue, initially thought to be exclusively responsible for analyzing auditory language comprehension, were responsible for analyzing musical syntax. (Maess et al., 2001)

Other studies investigating brain wave patterns, including the study by Miranda and Ullman (2007), found that the same region of the brain is responsible for memorizing and decoding vocabulary in a language and memorizing and decoding sequences of notes in music.

Thus, it is not surprising that researchers and educators have found a link between the use of songs and an increase in verbal memory (Abbott, 2002; Brewer, 2008; Jensen, 2000).

Researchers investigating first language (L1) learning through speech sequences compared to language learning through song sequences have found superior outcomes from the song-based stimuli. They also suggested that "the presence of pitch contours may enhance phonological discrimination since a change in pitch often accompanies syllable change" (Schön et al., 2008, p. 982). The use of songs with mild Alzheimer's patients revealed similar results. Patients learning words and sentence patterns through repeated song exposure showed better retention than those exposed to spoken lyrics (Moussard et al., 2012).

The increasing research on the advantages of using songs for first language (L1) acquisition has motivated second language acquisition (SLA) researchers to explore the potential of songs and other musical language forms in learning second languages (L2). Lindgren and Muñoz (2013) investigated factors that could predict EFL learners' English reading and listening performance in seven European contexts. Their findings suggested that out-of-school exposure to English through listening to a song, watching a film, reading and speaking in English, using the internet, and playing video games, was the second-best predictor of the participants' English reading and listening comprehension. Among the different sources of out-of-school exposure to English, they found that listening to songs was the most common source of foreign language exposure among the participants. Thus, the use of songs in L2 has the potential to foster language learning. However, further research is needed to operationalize and optimize the use of songs for second/foreign language acquisition.

1.1.2 How can language learning occur from songs?

Studies on language learning through reading and listening have revealed that various factors can impact the learning process. First is the number of encounters; learning is more likely to occur as encounters increase (Nation, 2015; Uchihara et al., 2019). This is especially true for learning the word forms that can become familiar through repeated encounters while their meaning stays elusive (Elgort & Warren, 2014). Second, the quality of context, a more informative context, can lead to higher levels of learning (Webb, 2008). Third, varied context; with every encounter, students can learn different aspects of word knowledge (e.g. Form-meaning connection, syntax, grammatical function, orthography, and association.), leading to deeper processing of target words (Webb, 2007). The fourth quality of encounter; as learners' level of attention to the target words increases, so do their learning gains (Nation, 2015). Fifth is context comprehensibility; learners need to understand the context in which they encounter the target word to learn specific aspects of word knowledge (i.e., Form-meaning connection) (Nation & Newton, 2009; Newton & Nation, 2021). Thus, if songs used for language learning fulfill these conditions, songs can be considered an auspicious source of language input.

Research is starting to show that songs have the potential to meet these requirements for language learning. For example, Tegge (2017) explored the lexical demand for popular songs by analyzing 1043 songs. Six hundred thirty-five were popular among language teachers, and 408 songs ranked most popular on the US Billboard charts. She found that to reach 95.1% lexical coverage of the popular songs on US billboard charts, learners needed to know the most frequent 3000-word families plus proper nouns. To reach 95.5%, lexical coverage of the songs selected by language teachers' students required knowledge of the most frequent 2000-word families plus

proper nouns. Thus, if language learners and teachers select songs appropriate for students' language level, songs can be a comprehensible source of language input.

Songs can provide large quantities of input for language learning, through which listeners will encounter the same words multiple times in different contexts (Murphey, 1992). Language learners commonly use songs as a source of language input (Kuppens, 2010; Lindgren & Muñoz, 2013). A study by Levey et al. (2011) investigating the listening habits of college students in North America found that, on average, the participants listened to 18.4 hours of music per week.

Songs are among the few sources of language input that pupils listen to multiple times without losing interest (Arnold & Herrick, 2017). A study by Pavia et al. (2019) exploring incidental vocabulary learning from listening to songs found a positive relationship between the number of times participants listened to a whole song and their performance on a vocabulary knowledge test. This study also indicated that students have the potential to learn different aspects of word knowledge through multiple exposures. For example, after listening to the same song five times, participants showed increased knowledge of collocations and spoken-form recognition. However, they found no significant results for word knowledge's spoken form meaning connection aspect. Thus, it is still unclear how songs can contribute to learning different aspects of vocabulary knowledge and how many encounters with target language items are needed for learning to occur.

Furthermore, conditions under which the participants were exposed to the target songs in previous studies (Pavia et al., 2019; Medina, 1993) were not ecologically valid. While learners listen to the same song multiple times, they tend to listen over a long time, with gaps between each listening session (Arnold & Herrick, 2017). Furthermore, the effect of time distribution on language learning gains is unclear. In a study by Serrano and Huang (2018) investigating

vocabulary learning from repeated reading and repeated reading while listening, they explored the effects of spaced distribution on learning gains. Their results suggested that spaced exposure can lead to superior long-term retention. Thus, research needs to investigate whether time distribution can influence learning gains from long exposure.

In addition, factors including the use of multimodal input, which combines the use of written text with auditory text, production activities such as singing, and the use of retrieval activities where learners recall what they have learned from memory, all of which have been used by teachers and learners to enhance the experience of song use for language learning (Arnold & Herrick, 2017) have yet to be examined empirically.

1.2 The Current Project

Through three individual studies, this thesis addresses some of the critical gaps in understanding how songs as a source of oral input can contribute to second/ foreign language learning. Vocabulary learning gains will be used as the measurement for language learning. Each study has a separate introduction, methodology, analysis, potential limitations, and implications.

1.2.1 Research Site

All studies in this project took place in the country of Thailand. The official language of this country is Thai; however, English is highly valued and became a mandatory part of the basic core curriculum of public education in 2008 (Ministry of Education Thailand, 2008). The students in this country share the same values and find English language education essential to their personal and professional development. However, students in Thailand struggle with learning English and have difficulty becoming proficient users of English. This is not surprising, as research shows that students learning English as a foreign language (EFL) do not have many

English language exposure opportunities outside the classroom (Webb & Nation, 2017). Thus, in the target context (i.e. Thailand), the student's English language knowledge remains solely attributed to what they have learned in school.

This project occurred inside a Thai school with the teachers' assistance and the school's principal. The school staff were kind enough to review the research materials, such as songs, prior to the start of the project to ensure no interference or overlap existed between the regular school curriculum and the research materials designed for this project.

Overall, the context of this research project (i.e. EFL) and the school staff's assistance increased the studies' internal validity. Thus, more robust conclusions may be drawn from the results, anticipating that the studies' outcomes were attributed to the studies' conditions rather than possible outside exposure.

1.2.2 Ethics

This research project used human participants for all three studies. Per the guidelines outlined by Canada's Tri-council research policies, ethics approval was obtained from the non-medical ethics board at Western University before launching the study (See Appendix A). Furthermore, consent was obtained from all individuals and institutions, directly and indirectly, involved in the research project.

Considering the site selected for this project (i.e., Thailand) prior to entering the country, approval was obtained from the *National Research Council of Thailand* (NRCT) (See Appendix B). This process involved submitting the university-approved research proposal, permission, and letter of support from the Thai school involved in the project.

The study's main participants were students learning English. Thus, upon arrival in Thailand, students and their guardians were presented with a bilingual information sheet outlining all necessary information about the study. All parties involved had the opportunity to review and complete written consent forms stating that their participation in the studies was voluntary and that they could withdraw from the studies at any time.

The study participants received a participant number upon completing the consent forms, which they used instead of indicating their names on any of the tests or forms completed for the study. The consent forms and all other information collected from the participants are stored at a safe location only accessible by the study's researchers. No personal information from the participant is released in any reports or publications that may result from this research project to protect the participants' confidentiality.

The school and the students participating in this study had the opportunity to learn new English vocabulary and a new method to learn and expose themselves to authentic English input. Nevertheless, a small gift as a token of appreciation was given to the study participants.

1.2.3 Procedure

The data collection portion of this project took a total of six weeks to complete. Before starting the study, the school and the teachers were presented with the information sheet and had the opportunity to ask questions and sign the written consent forms. Therefore, starting week one of the study, the work with the students who were the study's target participants began. The activities designed for this project were separate from the participant's Thai school curriculum and did not influence the topics or take away from what they studied with their Thai English teachers.

Before the start of the study, the school had randomly assigned students to each class; thus, this project took a quasi-experimental design. For the purposes of this project, each class was randomly assigned to one of the control or experimental groups for each study.

During week one, all participants were met twice. During the first session, all participants were presented with a bilingual information sheet, written consent forms, demographic information sheet, the bilingual vocabulary levels test for those in study one, and the updated vocabulary levels test for those in studies two and three. During session two, the participants in each study completed the specific vocabulary pretest designed for their specific study (i.e. Study 1, Study 2, & Study 3). Since all the vocabulary knowledge tests included at least a section on measuring formulaic sequences (FS) or, more specifically, collocation knowledge, an example of an FS that would have been familiar to the participants was written on the board and explained prior to starting the test. This example was taken from Pavia et al. (2019):

> (III) collocation recognition test:
> This example was written on the class board and explained prior to starting this section of the test:
> A) Sing a song B) Sing a food C) Sing a ball
> D) I don't remember any of these.

p. 754

Starting in week two, each study in this project followed a different procedure, discussed in detail in the following chapters.

1.3 References

Abbott, M. (2002). Using music to promote L2 learning among adult learners. *TESOL Journal, 11*, 10–17. https://doi.org/10.1002/j.1949-3533.2002.tb00061.x

Arnold, J. L., & Herrick, E. (2017). *New ways of teaching with music.* TESOL International Association. Alexandria, Virginia, USA: TESOL press.

Brandt, A., Gebrian, M., & Slevc, L. R. (2012). Music and early language acquisition. Frontiers in Psychology, 3, 327–327. https://doi.org/10.3389/fpsyg.2012.00327.

Brewer, C. B. (2008). *Soundtracks for learning: Using music in the classroom*. Bellingham, WA: LifeSounds Educational Services.

Coyle, Y., & Gomez Garcia, R. (2014). Using songs to enhance L2 vocabulary acquisition in preschool children. *ELT Journal*, *68*(3), 276–285. https://doi.org/10.1093/elt/ccu015

Elgort, I., & Warren, P. (2014). L2 vocabulary learning from reading: Explicit and tacit lexical knowledge and the role of learner and item variables. *Language Learning, 64*(2), 365–414. https://doi.org/10.1111/lang.12052

European Commission (2002). *Europeans' participation in cultural activities*. Retrieved from http://ec.europa.eu/public_opinion/archives/ebs/ebs_158_en.pdf.

Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: acquiring second language vocabulary through reading. *Reading in a Foreign Language, 11*(2), 207–223.

Jensen, E. (2000). *Music with the brain in mind: Enhance learning with music.* San Diego, CA: Corwin Press.

Kuppens, A. H. (2010). Incidental foreign language acquisition from media exposure. *Learning, Media and Technology*, *35*(1), 65–85. https://doi.org/10.1080/17439880903561876

Levey, S., Levey, T., & Fligar, B. J. (2011). Noise exposure estimates of urban MP3 player users. *Journal of Speech, Language, and Hearing Research, 54*, 263–277. https://doi.org/10.1044/1092-4388(2010/09-0283)

Levitin, D., & Menon, V. (2003). Musical structure is processed in "language" areas of the brain: A possible role for Brodmann Area 47 in temporal coherence. *NeuroImage, 20*, 2142–2152. https://doi.org/10.1016/j.neuroimage.2003.08.016

Lindgren, E., & Muñoz, C. (2013). The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism,10*, 105–129. https://doi.org/10.1080/14790718.2012.679275

Maess, B., Koelsch S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience, 4*(5), 540–545. https://doi.org/10.1038/87502

Medina, S. L. (1993). The effects of music on second language vocabulary acquisition. *National Network for Early Language Learning, 6*(3), 1-8.

Ministry of Education Thailand. (2008). *The Basic Core Curriculum.* Retrieved 17 November 2020 from http://www.act.ac.th/document/1741.pdf.

Miranda, R. A., & Ullman, M. T. (2007). Double dissociation between rules and memory in music: An event-related potential study. *NeuroImage, 38*, 331–345. https://doi.org/10.1016/j.neuroimage.2007.07.034

Moussard, A., Bigand, E., Belleville, S., & Peretz, I. (2012). Music as an aid to learning new verbal information in Alzheimer's Disease. *Music Perception: An Interdisciplinary Journal, 29*(5), 521-531. https://doi.org/10.1525/mp.2012.29.5.521

Murphey, T. (1992). The discourse of pop songs. TESOL Quarterly, 26, 770–774. doi:10.2307/3586887.

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, 20(2), 233–253. https://doi.org/10.2307/747758

Nation, I. S. P. (2015). Principles guiding vocabulary learning through extensive reading. *Reading in a Foreign Language, 27*, 136–145.

Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. New York: Routledge.

Newton, J., & Nation, I. S. P. (2021). Teaching ESL/EFL listening and speaking (Second edition.). Routledge, Taylor & Francis Group.

Pavia, N., Webb, S., & Faez, F. (2019). Incidental vocabulary learning through listening to songs. *Studies in Second Language Acquisition*, *41*(4), 745–768. https://doi.org/10.1017/S0272263119000020

Patel, A. (2008). *Music, language and the brain*. New York, NY: Oxford University Press.

Pellicer-Sanchez, A. (2017). Learning L2 collocations incidentally from reading. *Language Teaching Research, 21*(3), 381–402. https://doi.org/10.1177/1362168815618428

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical Learning in a Natural Language by 8-Month-Old Infants. Child Development, 80(3), 674–685. https://doi.org/10.1111/j.1467-8624.2009.01290.x

Randolph, P. T. (2017). Emotional involvement using music as the writing muse. In L. J. Arnold & E. Herrick. (2017). *New ways in teaching with music* (pp.35-36). TESOL International Association. Alexandria, Virginia, USA: TESOL press.

Richards, J. C. (2008). Second language teacher education today. *RELC. 39*(2), 158-177. https://doi.org/10.1177/0033688208092182

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition 70*, 27–52. https://doi.org/10.1016/S0010-0277(98)00075-4

Serrano, R., & Huang, H. (2018). Learning vocabulary through assisted repeated reading: How much time should there be between repetitions of the same text? *TESOL Quarterly, 52*(4), 971-994. https://doi.org/10.1002/tesq.445

Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition, 106*(2), 975–983. https://doi.org/10.1016/j.cognition.2007.03.005

Statistics Canada (1998). *Average time spent on activities, by sex.* Ottawa, Ontario. Retrieved 21 October 2007 from http://www40.statcan.ca/l01/cst01/famil36a.htm.

Tegge, F. (2017). The lexical coverage of popular songs in English language teaching. *System (Linköping)*, *67*, 87–98. https://doi.org/10.1016/j.system.2017.04.016

U.S. Department of Labor (2006). *American time use survey summary*. Washington, DC: U.S. Bureau of Labor Statistics. Retrieved 21 October 2007 from http://www.bls.gov/news.release/atus.nr0.htm.

Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language Learning, 69*(3), 559-599. https://doi.org/10.1111/lang.12343

Van Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System*, *41*, 609-624. doi: http://dx.doi.org/10.1016/j.system.2013.07.012

Vidal, K. (2003). Academic listening: A source of vocabulary acquisition? *Applied Linguistics, 24*, 56–89. doi:10.1093/applin/24.1.56. https://doi.org/10.1093/applin/24.1.56

Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning, 61*(1), 219-258. https://doi.org/10.1111/j.1467-9922.2010.00593.x

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics, 28*, 46–65. https://doi.org/10.1093/applin/aml048

Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language, 20,* 232–245.

Webb, S., & Chang, A. C. S. (2015a). How does prior word knowledge affect vocabulary learning progress in an extensive reading program*? Studies in Second Language Acquisition, 37,* 651–675. https://doi.org/10.1017/S0272263114000606

Webb, S., & Chang, A. C-S. (2015b). Second language vocabulary learning through extensive reading: How does frequency and distribution of occurrence affect learning*? Language Teaching Research, 19,* 667–686. https://doi.org/10.1177/1362168814559800

Webb, S., & Nation, P. (2017). *How vocabulary is learned.* Oxford, UK: Oxford University Press.

Chapter 2: Study 1

Incidental Vocabulary Learning through Massed Versus Spaced Listening to Songs: Replication and Extension

2.1 Abstract

Pavia, Webb, and Faez (2019) were among the first empirical studies to suggest that repeated listening to a song can promote knowledge of the spoken form of single words and collocations. While their results were statistically significant, the learning gains were small (i.e., 3.29% from one song and 8.67% from another song). One explanation for the small learning gains was the participants' limited prior vocabulary knowledge. It has been estimated that language learners need at least 95% lexical coverage of listening materials for vocabulary learning to occur (Van Zeeland & Schmitt, 2013). In Pavia et al. (2019), this condition was not met for many of the participants. Another explanation is that the participants listened repeatedly to the same song in a single session. Comparisons of massed versus spaced practice have tended to find an advantage for the latter (Kim & Webb, 2022). In the present study, 67 Thai EFL students with considerably better prior vocabulary knowledge than in the earlier study listened to the same song five times, either in a single session or spaced over five sessions. The results indicated learning gains of 26.89% in the massed learning condition and 21.31% in the spaced learning condition with a large effect size. These findings strengthen the argument for using songs in language learning and highlight the positive relationship between lexical coverage and conditions for unassisted incidental vocabulary learning.

Keywords: Replication and extension; Mass practice; Spaced practice; Vocabulary learning; Songs; Language acquisition

2.2 Introduction

Using songs for language acquisition in and outside of class has been common practice for decades. Teachers and learners find songs motivating and suggest that songs can help with pronunciation, grammar, and vocabulary learning (Arnold & Herric, 2017). However, very few studies have empirically investigated learning gains from songs. Medina (1993) investigated vocabulary learning by listening to children's storybook songs and compared the learning gains to those from exposure to the lyrics' spoken rendition. The results showed no significant differences between these conditions; however, the descriptive statistics showed higher learning gains among participants exposed to songs. Pavia et al. (2019) found similar learning patterns

from listening to songs. They compared vocabulary learning gains to those of a control group that received no exposure to the target words. Three experimental groups encountered the target words by listening to two songs. The experimental groups varied based on the number of times they listened to each song (one, three, or five times). On an immediate posttest, the experimental groups showed overall vocabulary gains of 3.29% from one song and 8.67% from another. Out of the three listening conditions, the group that encountered the songs more frequently (five times) showed higher learning gains than the other two listening groups (listening one and three times). Their vocabulary test measured learning of three aspects of word knowledge; spoken form recognition of single words, the form-meaning connection of single words, and spoken form collocation recognition. The experimental group participants outperformed the control groups on the spoken-form recognition of single words by 6.53% and the collocation recognition by 10.97%. These results were consistent with vocabulary learning gains from other sources of language input.

Listening to graded readers has shown gains of up to 16% for the spoken form of collocations (Webb & Chang, 2020). Van Zeeland and Schmitt (2013a) found that repeated encounters with target words in short listening passages fostered students' knowledge of spoken form and part of speech recognition. However, different aspects of word knowledge required more exposure than others. They concluded that knowledge of form-meaning connection would need more than 15 encounters before learners show any signs of knowledge. These reports suggest that listening to songs as a source of language input may be just as effective as using other sources of language input through listening, at least when the spoken form of words is concerned and perhaps their phraseological behaviour.

Pavia et al. (2019) provided a foundation for understanding the potential learning benefits of using songs for vocabulary acquisition. However, this study is one of the few studies that has examined vocabulary learning from songs and needs to be replicated. Therefore, the current study is built on the work of Pavia et al. (2019); it replicates the sections of their study that showed the highest levels of learning (i.e., listening five times) and used their research materials to test the generalizability of their results.

Furthermore, the previously mentioned study did not consider the effects of time distribution on vocabulary learning from songs. The time distribution effect refers to how massed (i.e., intensive) versus spaced (i.e., distributed) encounters with the source of language input can influence learning gains (Serrano & Huang, 2018). Rogers (2015) examined the effects of massed and distributed conditions on learning second language syntax and found no significant differences between the two groups on the immediate posttest. However, on the six-week delayed posttest, the participants in the distributed learning condition outperformed those in the massed learning condition. In another study, Suzuki (2017) examined the effects of massed (3.3-day interval) and spaced (7-day interval) learning of second language morphological structures for oral production in 60 Japanese students. The results indicated that the students in the massed learning condition did better on the seven-day and 28-day posttests with small to medium effect sizes. However, the massed learning condition in this study could still be considered spaced learning compared to the massed learning conditions described in previous studies using songs for language acquisition, where the learning occurred in only one session in one day. These contradictory findings warrant further investigation of the effect of time distribution on vocabulary learning. Thus, using the same research materials from Pavia et al. (2019), the current

study examined the impact of spaced encounters with songs on vocabulary learning gains and replicated the mass listening condition.

2.3 Literature Review

2.3.1 Incidental Vocabulary Learning and Contributing Factors

Incidental language learning and, more specifically, incidental vocabulary learning have received various definitions over the years (Hulstijn, 2003). In the present study, incidental vocabulary learning is defined as learning vocabulary items and patterns throughout an activity (i.e., listening to a song) where the learners primarily engage with the activity for enjoyment and understanding the message behind the words (Boers, 2021). Thus, vocabulary learning occurs as a by-product of an activity (Webb, 2020). Incidental vocabulary learning can be essential to improving the depth and breadth of vocabulary knowledge. Studies looking at the vocabulary knowledge needed to understand written and spoken text in English without support estimate that language users need to know knowledge of the most frequent 3,000-word families for conversation (Nation, 2006; Van Zeeland & Schmitt, 2013), 4,000-word families for academic spoken text (Dang & Webb, 2014), and at least 8,000-word families to understand written texts (Nation, 2006).

Given the large number of words needed, learning through deliberate study is unrealistic. In fact, Webb and Chang (2012a) found that adult Taiwanese students struggle to master the knowledge of the most frequent 3,000-word families after years of mainly deliberate English language instructions. This is far below the number of word families needed for unassisted use of English and does not come close to the vocabulary size of L1 English, educated adults with a vocabulary size of approximately 17,000-word families (Goulden et al., 1990). Furthermore, a meta-analysis investigating the efficacy of intentional vocabulary learning found that the initial

large learning gains (58%-60%) suggested by these studies are not retained over a long period of time, and true learning gains (25%-39%) from deliberate instructions are considerably less than what was previously believed (Webb, Yanagisawa, & Uchihara, 2020). Thus, deliberate study alone is not enough, and learners need to engage in activities that allow incidental vocabulary learning (Webb, 2020).

In the last three decades, studies have explored incidental vocabulary learning from various sources of input, including reading, listening, reading while listening, and viewing (Hulstijn et al., 1996; Pavia et al., 2019; Pellicer-Sánchez, 2016; 2017; Peters & Webb, 2018; Reynolds, 2015; Rodgers & Webb, 2020; Teng, 2018; van Zeeland, & Schmitt, 2013). The results proposed several factors that could influence the incidental acquisition of language items. First, repetition and frequency of encounters; As early as the 1970's researchers have advocated for repeated encounters of target language items for learning to occur. Saragi et al. (1978) suggested that vocabulary learning can occur through repeated encounters with items through extensive reading exercises and that repetition is necessary for L2 learners and L1 users of English. In their study, L1 English users needed more than ten encounters with target words for learning to occur. A meta-analysis of correlational studies by Uchihara et al. (2019) found a medium effect (r= .34) for repetition on incidental vocabulary learning. However, the number of encounters needed for learning to occur was different across studies. One factor explaining this variance is the aspect of word knowledge each study used to measure learning gains. For example, after one encounter in reading activities, Chen and Truscott (2010) found learning gains for 43% of the target words, and Webb (2007) found 67% learning gains for the form recognition aspect of word knowledge. In contrast, when using meaning recognition to measure learning gains, studies have found anywhere between two encounters (Rott, 1999), two to four (Pellicer-

Sánchez & Schmitt, 2010), eight to ten (Waring & Takaki, 2003), and more than ten encounters (Pigada & Schmitt, 2006) needed for learning to occur. In addition, these studies suggested that even more encounters are needed for recall knowledge, with seven encounters for form recall in Chen and Truscott (2010), ten encounters for meaning recall in Webb (2007) and between 10-17 encounters for meaning recall in Pellicer-Sánchez and Schmitt (2010). A recent study by González-Fernández and Schmitt (2020) found that different word knowledge components are acquired at different times, and there is a consistent pattern of acquisition. Their results suggested that receptive knowledge is acquired first, as the participants started recognizing the link between form and meaning. Next, knowledge of collocate form recognition and recall knowledge, such as collocate form recall, is acquired. Lastly, multiple-meaning recall is acquired, which tends to be a more difficult component of word knowledge (González-Fernández & Schmitt, 2020, p.493). These findings can explain the findings of Pavia et al. (2019) to a degree. Pavia et al. (2019) found that form recognition knowledge of words was developed from listening to the first song, and collocation recognition knowledge was developed from listening to the second song. In addition, form-meaning recognition showed no learning gains from either song, suggesting that different aspects of word knowledge develop at different times also when using songs as the source of input. However, the 6.53% (0.52 words) learning gains in spoken-form recognition and 10.97% (0.76 collocates) for collocation recognitions were much smaller than in previous studies. For example, Webb and Chang (2018) found 16% (2.58 words) learning gains for the spoken form of collocations. Thus, further research is needed to understand how learning can occur from incidental listening to songs and how the frequency of exposure can affect the different aspects of word knowledge from songs.

Frequent exposure to target language items can transpire in different ways. Repeated exposure through repeated listening or reading the same spoken/written text is found to be essential for learning a second language (DeKeyser, 2015; Horst & Meara, 1999; Webb & Chang, 2012b). This is not surprising as researchers have found that learners have limited processing capacity, meaning if they focus on form processing, they will have difficulty processing the semantic aspects of target language items (Newton, 2020). In the *Type of Processing-Resource Allocation* (TOPRA) framework, Barcroft (2002; 2013) explains that learners struggle with limited attention capacity, and if they allocate attention to one aspect of input, it may reduce the cognitive resources available to process the other aspects of input. Furthermore, he identifies repeated exposure as one of the key components that can benefit word learning (Barcroft, 2013). Pavia et al. (2019) found listening five times to the same song to be the most effective condition for vocabulary learning in their study. However, they only tested repeated listening without spacing (massed learning). Research in cognitive psychology (Cepeda et al., 2006) and second language acquisition (Kim & Webb, 2022) suggest that spacing can benefit learners' long-term consolidation of knowledge. However, the majority of the studies investigating the effects of spaced practice have focused on deliberate learning of target language items (Bird, 2010), and only a few have considered spaced practice with incidental learning (Uchihara et al., 2019; Kim & Webb, 2022). While it is well established that spaced practice results in more durable learning in deliberate learning conditions, mixed results have been found for incidental learning conditions. Uchihara et al. (2019) meta-analysis exploring the predictability of different factors on incidental vocabulary learning used the spacing effect as a moderator. Their results suggested that there was a larger learning effect for repeated exposure to the target language in massed conditions (r = 0.38, 95% CI [0.31, 0.45]) compared to spaced

conditions (r = .23, 95% CI [.12, .34]). Another study by Serrano and Huang (2023) explored the effects of different schedules on deliberate and incidental vocabulary learning through reading. Their results indicated the short-spaced repeated reading had a significantly more positive effect on both deliberate and incidental learning conditions. However, their short-spaced intervention intervals were one day apart, which is more spaced than the mass listening condition used by Pavia et al. (2019). Thus, the second factor influencing incidental acquisition of language items that need to be further explored is the optimal schedule for repeated listening to a song comparing massed and spaced learning conditions.

Finally, prior word knowledge can affect language learning progress (Horst et al., 1998; Zahar et al., 2001). Webb and Chang (2015a) found that learners' prior vocabulary knowledge greatly impacted learning gains from a 37-week extensive reading program. They found individuals in the high-level proficiency group to gain approximately 35% more than those in the low-level proficiency group on their immediate posttest of vocabulary knowledge. On the delayed posttest high-level proficiency group once again outperformed the low-level group while showing a regression of 4.17%, which was less than the 7.07% regression observed in the low-level group. More proficient learners with larger vocabulary sizes tend to have more extensive lexical coverage of the input, making it easier for them to understand and learn the remaining unknown words from context (van Zealand & Schmitt 2013b). The number of studies examining incidental vocabulary learning from songs is very limited; therefore, it is not clear how proficiency level and lexical coverage affect vocabulary learning from songs, so more research is needed to understand this relationship better.

2.3.2 Use of Songs for Language Acquisition

Songs continue to be the number one source of exposure to second/foreign language input outside of the classroom (Lai et al., 2015; Peters, 2020). However, reports of students' language proficiency and its relationship with listening to songs have failed to find a connection between language acquisition and song use (Peters, 2020). The question of how songs can be effectively used for ESL/EFL development remains.

Most studies investigating different aspects of language learning through songs have used songs as the medium for the deliberate teaching of language items. For example, Tegge (2015) examined verbatim recall and recognition of target text in 105 participants from Serbia, Germany, and Belgium. They used a series of activities, including pre-teaching of target words, listening, reading, singing, and completing gap-fill exercises using songs, poems, and prose texts as the sources of language input. The results demonstrated that songs compared to the other two types of texts for language learning, resulted in higher learning gains with a large effect size. In another study, Roohani and Akbarpour (2016) measured 100 Iranian EFL students' vocabulary learning by deliberately teaching target vocabulary through flashcards and pictures while using songs or a simple non-musical story to practice target vocabulary items. Their descriptive statistics suggested higher learning for individuals in the song group. However, these findings were not statistically significant.

Mannarelli and Serrano (2022) examined receptive and productive vocabulary learning from implicit and explicit instructions using songs. In their study, 27 participants were divided into two groups, explicit instruction, where the participants deliberately focused on the target words and implicit condition, where the participants focused on words in the songs that were not used for vocabulary testing in a pre-post-delayed test format. They found significant relative

gains of 28.72% for the implicit group and 47.82 % for the explicit group. While these results seem promising, words of caution are in order here: lack of a control group, the small sample size, and the likelihood that instructing the implicit group to focus on words other than those on the vocabulary test reduced the attention they might otherwise have given to the latter.

Overall, these findings can be taken as encouragement for language teachers to use songs in addition to other types of text in their classrooms. However, they overlooked potentially the most important affordance of song use for language development. As mentioned earlier, songs are the most common source of language input among ESL/EFL learners outside the classrooms. More research is therefore needed to help optimize *incidental* language acquisition from songs without taking away from the natural experience.

2.3.3 The Need for Replication and Extension Studies

In addition to Pavia et al. (2019), several studies have explored incidental vocabulary learning from songs. However, several methodological issues and missing information in the treatment conditions make it difficult to interpret the results of these studies and make them irreplicable. Replicability is a fundamental aspect of the scientific method. Porte and McManus (2019) point out, “No one piece of experimental research (or researcher!) can include, or control for, all the many variables that might affect the outcome.” (p.4). Replication studies can check the instruments’ reliability and increase the results’ generalizability when tested with different participants across different contexts (Marsden, 2020). There are few replication studies investigating incidental vocabulary learning from songs to date. Tilwani et al. (2022) resemble a partial replication of previous research (albeit without explicitly acknowledging it) because they used the same song that was used in the study by Bahrami et al. (2019) and the same three aspects of word knowledge (i.e., spoken-form recognition, form-meaning connection, and

collocation recognition) were tested as in the study by Pavia et al. (2019). However, Tilwani et al. (2022) failed to provide items from their vocabulary knowledge test to allow full replicability and to help readers compare their results to the previously mentioned studies.

Marsden et al. (2018) found a positive association between the extent to which researchers provided their research materials and the likelihood that replication studies support the original research outcomes. This highlights the importance of providing a detailed methodology section in study reports and of providing the instruments used for research. However, a large portion of studies published lack a detailed methodology section. Poor availability of materials reduces researchers' capacity to compare results and check the reliability of previously collected data. For example, Köksal et al. (2013), measuring "*The effects of music on achievement, attitude, and retention in primary school English lesson,"* includes a single short paragraph outlining the methodology section of their study, which leaves out crucial information. They mention that the students in the experimental group learned vocabulary items through songs while the students in the control group "were taught the same new words with the methods in the current education program." (p.1898) without providing any further detail regarding those methods. They do not explain how the participants were exposed to the target words in the song condition and what the "current education program" involves, making this study non-replicable and the reliability of the results questionable.

The lack of replication and replicable studies investigating incidental vocabulary learning from songs warrants replication studies in this area. Furthermore, as mentioned earlier, many unknown factors, such as the effect of distributed exposure, still need to be examined. Therefore, a replication and extension study is needed at this time. Replication extension studies provide further evidence regarding the topic being investigated by combining the results from prior

studies, allowing for a more precise account of effect sizes and increasing the generalizability of the results (Bonett, 2012).

2.4 Current Study

The current study investigated incidental vocabulary learning through listening to songs by replicating the most effective condition of Pavia et al. (2019), listening five times to a single song. Vocabulary knowledge was measured by comparing scores between the pre-test to the immediate posttest and delayed posttest on single-word form recognition, form-meaning connection, and collocation recognition, using the tests designed by Pavia et al. (2019). In addition to the replication portion, a new variable, the effect of spacing on vocabulary learning, is introduced in the extension portion of this study.

2.5 Research Questions

2.5.1 Does massed and spaced listening to songs contribute to L2 incidental vocabulary learning? (Adapted from Pavia et al., 2019)

1. A. Word knowledge is broken into three aspects (Adapted from Pavia et al., 2019)
    I. Spoken-form recognition
    II. Form-meaning connection
    III. Collocation recognition

2.5.2 How does input distribution (massed versus spaced) affect receptive vocabulary learning gains from listening to songs?

2.A. Word knowledge is broken into three aspects (Adapted from Pavia et al., 2019)
    I. Spoken-form recognition
    II. Form-meaning connection

III. Collocation recognition

2.6 Method

2.6.1 Participants

The participants were 67 Thai-speaking students in Matthayom six (Grade 12), which is the last year of Thai high school. Their age ranged from 17 to 19 (M = 17.79, SD = 0.51). English is a compulsory subject in most Thai schools, and it is introduced to students as early as kindergarten or preschool at the age of four or five. Most of the participants in this study received formal English education for approximately 13 years. Their regular English teacher placed them at the beginner-intermediate proficiency level, and they were familiar with the first 1000 most frequent word families (M = 26.62 out of 30, SD = 3.5) and were learning the most frequent 2000 word families (M = 23.17 out of 30, SD = 5.5) as measured by the bilingual (Thai -English) vocabulary levels test (VLT) from Nation's website (https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/vocabulary-tests) (See Appendix C). This was the test used by Pavia et al. (2019); thus, the results are comparable to the results of the previously mentioned study. Based on Webb and Chang (2015a), a correct response on the VLT represents knowledge of 33.3 words; thus, the participants in this study had knowledge of approximately 1,660 of the most frequent 2,000-word families.

The participants were randomly assigned to one of four classes by the Thai school. Each intact class was then randomly assigned to one of the four conditions in this study. Their VLT scores were normally distributed, as indicated by the insignificant ($p > 0.05$) results of the Kolmogorov-Smirnov test of normality.

2.6.2 Research Instruments

This study used the research instruments from Pavia et al. (2019).

2.6.2.1 Song

Of the two songs used in the previous study, only one, "Die a happy man" by Thomas Rhett (2015), was used. The lyrics of the song are available at "https://www.azlyrics.com/lyrics/thomasrhett/dieahappyman.html" (See Appendix E). Using Pavia et al.'s justification for song selection, this song is an appropriate choice for this study's participants. First, based on communication with other students in grade 12 and the teachers in Thailand, this song is age-appropriate and exciting for students in this age group. This song uses no vulgar language, and it is new to the participants of this study.

The lyrics of the song are analyzed by the same software used in Pavia et al. (2019), entitled “Range” (Heatley & Nation, 2002) and Nation's (2017) British National Corpus/ Corpus of Contemporary American English (BNC/COCA) word family lists (p.7). The analysis (see Table 2.1) was found to be the same as those in the previous study. It indicates that the participant would need to know the most frequent 2K word families to reach 95% lexical coverage of the song. As mentioned above, the participants in this study knew approximately 1,660 of the most frequent word families and therefore fell slightly below the proposed 95% threshold (Van Zeeland & Schmitt, 2013a). The participants in Pavia et al. (2019) had much poorer scores on the VLT: They were, on average, familiar with only 430 high-frequency word families. The participants' substantially better prior vocabulary knowledge in the present study can thus give insight into the effects of greater prior knowledge and proficiency on learning gains when compared to the previous study.

Table 2.1 Lexical frequency Profile: "Die A Happy Man"

Level	Token%	Cumulative coverage %
1	90.51	90.51
2	3.48	93.99
4	0.32	94.31
5	0.63	94.94
6	0.32	95.26
31	1.90	97.96
32	1.90	99.06
33	0.32	99.38
Not on the list	0.63	100.01

Similar information is found in Pavia et al. (2019, p.8)

2.6.2.2 Target words

The song includes 19 single-word items (See Table 2.2) and seven collocations. The previous study used mutual information (MI) scores of above 3 for selecting collocations. In addition to MI scores retrieved from the COCA (1990-2019) corpus, the face validity of target collocations was tested in this study. Twelve dominant English speakers were asked to identify collocations in the target song. As a result, the same target collocations were selected (see Table 2.3). However, there is a limitation with two of the target collocations, "bottle" and "wine," "listen" and "radio" if the students can recognize the meaning of each of the single words in the pairs, they may be able to recognize that the target words go together. Nevertheless, if the students know the meaning of these words, they would be able to recognize the correct pairs on the pre-test and the posttests and show no learning gains. Therefore, these items were not removed as they did not influence the conclusions drawn from the study.

Table 2.2 Target Single Words

Song	Target words	Frequency (*fr*)
Die a Happy Man	Die	1000
	Enough	1000
	Between	1000
	Build	1000
	Last	1000
	True	1000
	Under	1000
	Fancy	2000
	Escape	2000
	Coast	2000
	Dark	2000
	Star	2000
	Knee	2000
	Saint	3000
	Fireplace	5000
	Destination	5000
	Mansion	5000
	Vacation	6000
	Masterpiece	7000

Similar information is found in Pavia et al. (2019 p.9)

Table 2.3 Target Collocations

Song	Target Collocations	*fr*	MI Score
Die a Happy Man	Bottle Wine	2312	7.80
	Pouring Rain	584	7.80
	Northern Lights	546	5.28
	Wildest Dreams	774	10.69
	Listen Radio	749	3.44
	Sports Car	1525	3.78
	No Doubt	30336	4.76

Note "fr" indicates how frequently the collocates occur.*

Similar information is found in Pavia et al. (2019, p.9)

2.6.2.3 Dependent Measures

The test scores from a multiple-choice test used by Pavia et al. (2019) was used in this replication (see Appendix F). The test includes three sections, each measuring a different aspect of vocabulary knowledge. The first two sections evaluate receptive knowledge of single words by measuring spoken-form recognition and form-meaning connection. The last section evaluates receptive knowledge of multiword sequences by measuring learning gains through collocation recognition. The students were exposed to the test questions orally from a recording and were asked to select the correct option on the multiple-choice in pencil and paper format.

2.6.3 Procedure

The four student groups were assigned to the massed or the spaced listening sections. For each section, one group was the control group and the other the experimental group. Chapter one described the procedure for collecting consent forms, demographic information forms, bilingual VLT, and the vocabulary knowledge test, which was implemented during week one.

During week two, control group one (C1) completed the immediate posttest without exposure to the target song. Experimental group one (E1) listened to the target song five times, followed by the immediate posttest. For the duration of the study, if the participants asked any questions that could have compromised the study results, they were told that all their questions would be answered at the end of our last session together. These questions included asking about the name of the song or asking about the meaning of a particular word in the song. After a two-week delay, E1 and C1 completed the delayed posttest in week four. These procedures replicated the mass listening condition that showed the highest learning gains in Pavia et al. (2019).

From weeks two to four, experimental group two (E2) listened to the target song once every 4.5 days and completed the immediate posttest after the final listening session in week

four. Control group two (C2) was not exposed to the target song from weeks two to four. They completed the immediate posttest in week four. Both E2 and C2 completed the two weeks delayed posttest in week six. This section measured the effects of spaced listening on vocabulary knowledge. Table 2.4 provides an overview of the procedures.

2.6.4 Analysis

All data collected from this study was numeric; thus, quantitative statistical analyses were performed (Mujis, 2011). *IBM SPSS Statistics* version 28 data analysis software generated descriptive statistics such as means, standard deviations, etc., and inferential statistics, including the multivariate analysis of covariance variance (MANCOVA), comparing the results between the pretest, posttest, and delayed posttest within and between groups. Considering that the Thai school randomly assigned students to different groups, the Kolmogorov-Smirnov normality test was used to ensure that the participants were normally distributed between groups based on the VLT test. Mauchly's Sphericity test was used to assume homogeneity of within-group variance. If the assumption of sphericity had been violated, the Greenhouse-Geisser was used to adjust the *df* (Warner, 2013).

Table 2.4 Timeline

Group	Week1	Week 2		Week 3	Week 4		Week 5	Week 6
Mass listening sections								
	Session 1	Session 2		Session 3	Session 4		-	-
Control (C1)	• Consent Form • Demographic information form • VLT • Pre-test	Immediate posttest		Regular Thai school curriculum	Delayed posttest		Regular Thai school curriculum	Regular Thai school curriculum
Experimental (E1)		• Listen to target song 5X • Immediate posttest		Regular Thai school curriculum	Delayed posttest		Regular Thai school curriculum	Regular Thai school curriculum
Spaced listening sections								
	Session 1	Session 2	Session 3	Session 4	Session 5	Session 6	Session 7	Session 8
Control (C2)	• Consent Form • Demographic information form • VLT Pre-test	Regular Thai school curriculum	Regular Thai school curriculum	Regular Thai school curriculum	Regular Thai school curriculum	Immediate posttest	Regular Thai school curriculum	Delayed posttest
Experimental (E2)		Listen to target song once	Listen to target song once	Listen to target song once	Listen to target song once	Listen to target song once followed by immediate posttest	Regular Thai school curriculum	Delayed posttest

Note. The treatment sessions for E2 were 4.5 days (i.e., four to five days) apart.

2.7 Results

The descriptive statistics based on the participant's scores on the vocabulary knowledge test are presented in Tables 2.5 and 2.6. The maximum possible score on the spoken-form recognition and form-meaning connection was 19, and seven on the collocation recognition. These equal the maximum score of 45 for all aspects of word knowledge. Repeated measures ANOVA was used to answer the first research question investigating the interaction between time and exposure to a song under massed and spaced listening conditions.

2.7.1.1 Overall vocabulary learning from massed and spaced listening

In the massed listening sections, the analysis showed that the assumption of sphericity had been violated for the control (C1) ($\chi^2(2) = 14.758$, $p < 0.001$) and the experimental (E1) ($\chi^2(2) = 28.577$, $p < 0.001$) groups; therefore, a Greenhouse-Geisser correction was used. The within-participant main effect (time) indicated no statistically significant F (1.230, 19.678) = 0.008, $p = 0.992$, results for C1. In contrast, the results for E1 indicated that the within-participant main effect (time) was statistically significant, F (1.114, 21.162) = 49.920, $p < 0.001$ with partial $\eta^2 = 0.724$ (Very large effect size).

In the spaced listening sections, the analysis showed that the assumption of sphericity had been met for the control (C2) ($\chi^2(2) = 5.183$, $p = 0.075$) group. However, this assumption was violated for the experimental (E2) ($\chi^2(2) = 33.549$, $p < 0.001$) group, and the Greenhouse-Geisser correction was used. The within-participant main effect (time) indicated no statistically significant results for C2: F (2, 24) = 0.883, $p = 0.427$. In contrast, the results for E2 indicated that the within-participant main effect (time) was statistically significant, F (1.056, 16.903) = 55.948, $p < 0.001$ with partial $\eta^2 = 0.778$ (Very large effect size) (see Figure 2.1).

Table 2.5 Descriptive statistics for overall scores on the vocabulary knowledge test: Means (and SDs)

Participant subgroups	All Aspects of Word Knowledge		
	Pretest	Immediate posttest	Delayed posttest
C Mass (C1) (*n= 17)*	10.82 (2.42)	10.82 (3.70)	10.76 (2.93)
E Mass (E1) (*n= 20)*	21.65 (5.76)	33.75 (5.81)	31.60 (5.37)
C Spaced (C2) (*n = 13)*	15.84 (2.82)	15.62 (2.14)	15.23 (2.31)
E Spaced (E2) (*n= 17)*	17.06 (4.80)	25.88 (4.53)	25.06 (4.34)

Note. The maximum score was 45. The total number of participants was N= 67.

Table 2.6 Descriptive statistics of different aspects of word knowledge: Means (and SDs)

Participant subgroups	Spoken-form recognition			Form-meaning connection			Collocation recognition		
	Pretest	Immediate posttest	Delayed posttest	Pretest	Immediate posttest	Delayed posttest	Pretest	Immediate posttest	Delayed posttest
C Mass (C1) (*n= 17)*	4.71 (1.40)	4.88 (2.73)	4.71 (1.96)	4.65 (1.41)	4.24 (1.39)	4.41 (1.50)	1.47 (0.72)	1.71 (0.919)	1.64 (0.70)
E Mass (E1) (*n=20)*	8.65 (3.06)	14.65 (2.21)	13.65 (2.06)	10.95(3.79)	14.70 (5.23)	13.90 (4.85)	2.05 (1.19)	4.40 (1.23)	4.05 (0.99)
C Spaced (C2)(*n =13)*	7.46 (1.45)	7.08 (0.95)	6.92 (1.11)	7.08 (1.49)	7.15 (1.62)	6.92 (1.12)	1.31 (0.63)	1.38 (0.51)	1.38 (0.65)
E Spaced (E2) (*n= 17)*	7.88 (2.34)	12.35 (5.76)	11.94 (2.53)	7.71 (2.28)	9.65 (1.41)	9.35 (1.32)	1.47 (1.37)	3.88 (2.03)	3.76 (1.99)

Note. The maximum score on the spoken-form recognition and the form-meaning connection were 19, and the maximum score on collocation recognition was 7.

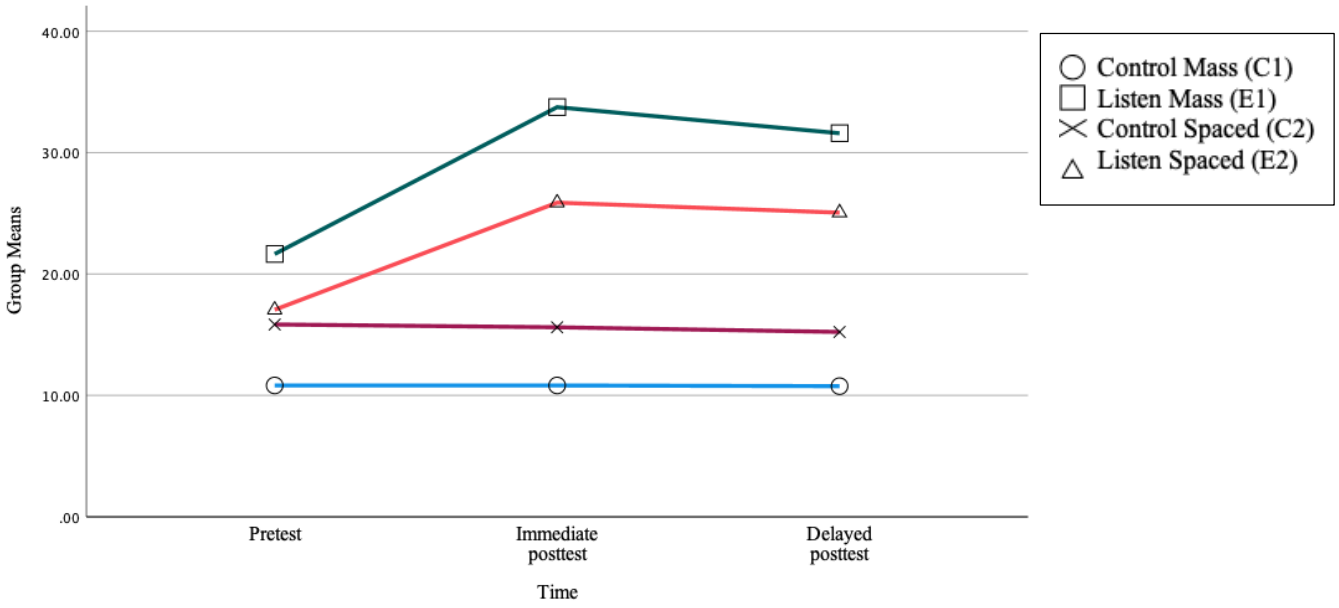


Figure 2.1. Group means for overall vocabulary learning over time

2.7.1.2 Learning of different aspects of word knowledge from massed and spaced listening

2.7.1.2.1 Spoken-form Recognition

In the massed listening sections, the analysis showed that the assumption of sphericity had been violated for C1 ($\chi^2(2) = 23.794$, $p < 0.001$) and E1($\chi^2(2) = 25.471$, $p < 0.001$); therefore, a Greenhouse-Geisser correction was used. The within-participant main effect (time) indicated no statistically significant results for C1: F (1.114, 17.824) = 0.147, $p = 0.733$. In contrast, the results for E1 indicated that the within-participant main effect (time) was statistically significant: F(1.138, 21.627) = 46.378, $p < 0.001$ with partial $\eta^2 = 0.709$ (Very large effect size).

In the spaced listening sections, the analysis showed that the assumption of sphericity had been met for C2 ($\chi^2(2) = 2.895$, $p = 0.235$). However, this assumption was violated for E2 ($\chi^2(2) = 29.116$, $p < 0.001$), and the Greenhouse-Geisser correction was used. The within-participant main effect (time) indicated no statistically significant results for C2: F (2, 24) = 1.636, p =

0.216. In contrast, the results for E2 indicated that the within-participant main effect (time) was statistically significant: $F(1.077, 17.237) = 44.62$, $p < 0.001$ with partial $\eta^2 = 0.736$ (Very large effect size) (see Figure 2.2).

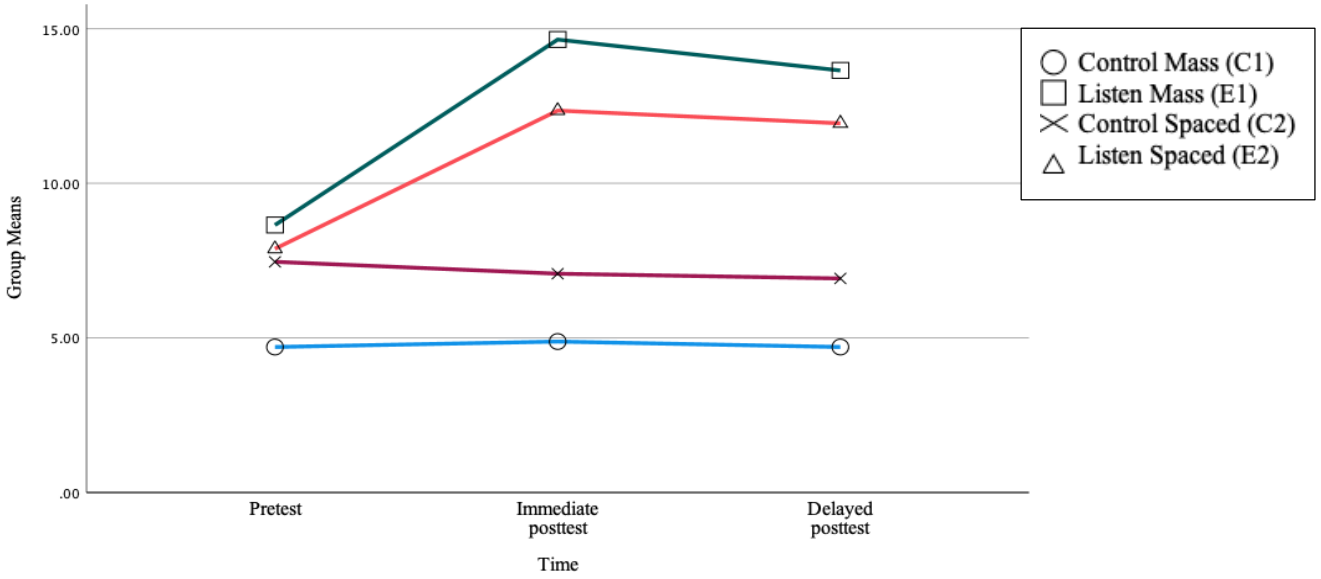


Figure 2.2. Group means for spoken-form recognition over time

2.7.1.2.2 Form-meaning Connection

In the massed listening sections, the analysis showed that the assumption of sphericity had been violated for C1 ($\chi^2(2) = 12.91$, $p = 0.002$) and E1 ($\chi^2(2) = 27.183$, $p < 0.001$); therefore, a Greenhouse-Geisser correction was used. The within-participant main effect (time) indicated no statistically significant ($F(1.27, 20.29) = 1.252$, $p = 0.289$) results for C1. In contrast, the results for E1 indicated that the within-participant main effect (time) was statistically significant ($F(1.124, 21.359) = 16.412$, $p < 0.001$ with partial $\eta^2 = 0.463$) with a very large effect size.

In the spaced listening sections, the analysis showed that the assumption of sphericity had been met for C2 ($\chi^2(2) = 2.709$, $p = 0.258$). However, this assumption was violated for E2 ($\chi^2(2) = 34.872$, $p < 0.001$), and so the Greenhouse-Geisser correction was used. The within-participant

main effect (time) indicated no statistically significant (F (2, 24) = 0.332, p = 0.721) results for C2. In contrast, the results for E2 indicated that the within-participant main effect (time) was statistically significant (F (1.051, 16.823) = 11.003, $p < 0.001$ with partial $\eta^2 = 0.820$) with a very large effect size (see Figure 2.3).

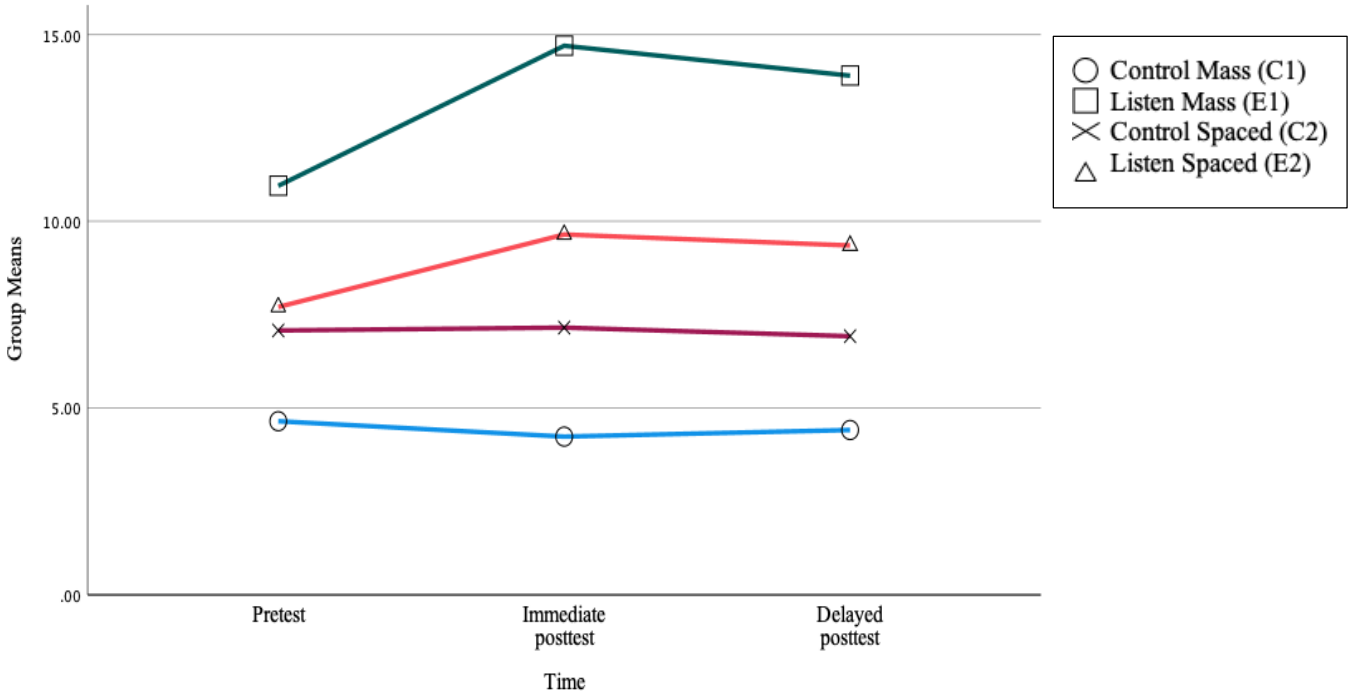


Figure 2.3. Group means for form-meaning connection over time

2.7.1.2.3 Collocation Recognition

In the massed listening sections, the analysis showed that the assumption of sphericity had been met for C1 ($\chi^2(2) = 1.69$, $p > 0.05$). However, this assumption was violated for E1 ($\chi^2(2) = 26.505$, $p < 0.001$), and the Greenhouse-Geisser correction was used. The within-participant main effect (time) indicated no statistically significant (F (2, 32) = 1.089, p = 0.349) results for C1. In contrast, the results for E1 indicated that the within-participant main effect (time) was statistically significant (F (1.130, 21.461) = 33.594, $p < 0.001$ with partial $\eta^2 = 0.639$) with a very large effect size (see Figure 2.3).

In the spaced listening sections, the analysis showed that the assumption of sphericity had been violated for C1 ($\chi^2(2) = 6.193$, $p = 0.045$) and E1 ($\chi^2(2) = 21.486$, $p < 0.001$); therefore, a Greenhouse-Geisser correction was used. The within-participant main effect (time) indicated no statistically significant (F (1.398, 16.77) = 0.093, $p = 0.844$) results for C1. In contrast, the results for E1 indicated that the within-participant main effect (time) was statistically significant (F (1.136, 18.169) = 72.864, $p < 0.001$ with partial $\eta^2 = 0.820$) with a very large effect size.

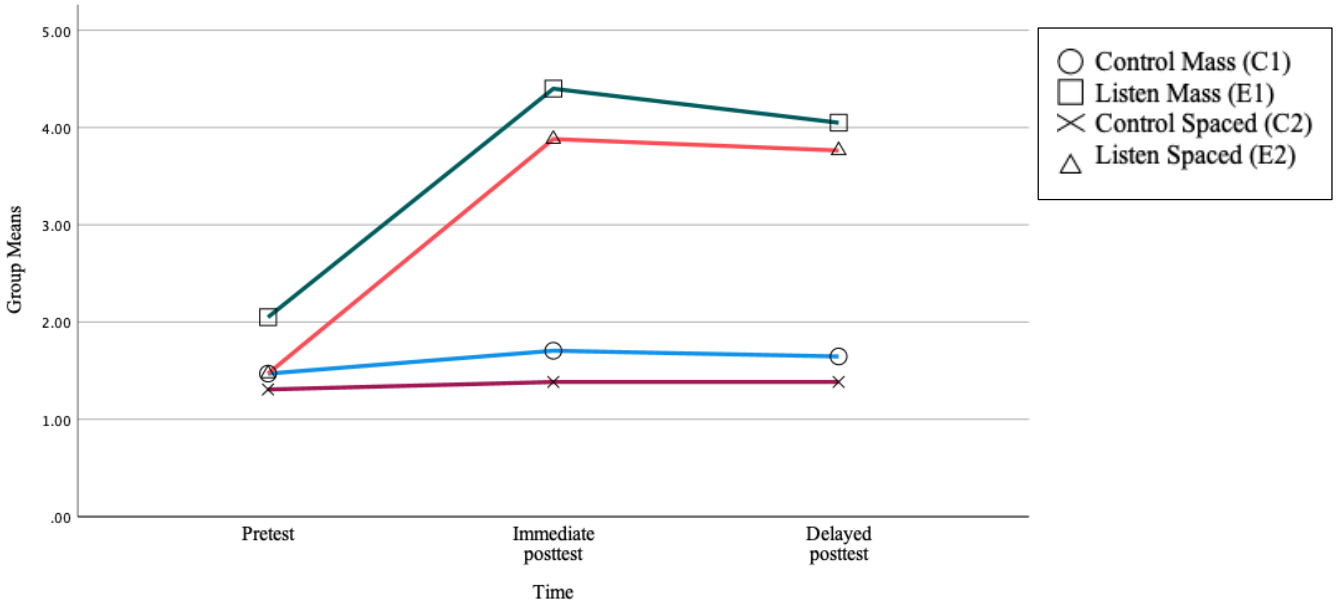


Figure 2.4. Group means for collocation recognition over time

2.7.1.2.4 Vocabulary Learning gains

The pairwise comparison between test times (pretest, immediate posttest, and delayed posttest) was used to determine the vocabulary learning gains for each aspect of word knowledge under massed and spaced listening distribution conditions.

The results indicated no significant gains in any aspects of word knowledge for the control groups (C1 and C2). The participants in the experimental group (E1) showed significant increases in their scores from the pretest to the immediate posttest with a large effect size

(*d*=2.247) and from the pretest to the delayed posttest with a large effect size (*d*=1.917) for the spoken-form recognition. Participants recognized an additional six words on average on the immediate posttest and an additional five words on the delayed posttest compared to the pretest under the massed listening condition.

Under the spaced listening condition (E2), participants also showed a significant increase in their scores from the pretest to the immediate posttest with a large effect size (*d*=1.017) and from the pretest to the delayed posttest with a large effect size (*d*=1.660) for the spoken-form recognition aspect of word knowledge. They recognized 4.471 more words on average on the immediate posttest and an additional 4.059 words on the delayed posttest compared to the pretest.

On the form-meaning connection aspect of word knowledge, results showed a significant increase from the pretest to the immediate posttest (*d*=0.821) and from the pretest to the delayed posttest (*d*=0.678) with large and medium effect sizes under the massed listening conditions (E1). Participants could recognize an additional 3.75 words on the immediate posttest and additional 2.95 words on the delayed posttest compared to the pretest.

The participants in the spaced learning conditions (E2) showed a significant increase in their scores on the form-meaning connection aspect of the tests with a large effect size for both the pretest to immediate posttest (*d*=1.023) and the pretest to delayed posttest (*d*=0.880). They could recognize the meaning of approximately 1.941 additional words on the immediate posttest and approximately 1.647 more words on the delayed posttest compared to the pretest.

On the final aspect of word knowledge, collocation recognition, participants showed a significant increase from the pretest to the immediate posttest (*d*=1.942) and from the pretest to the delayed posttest (*d*=1.827) with large effect sizes under the massed listening conditions (E1).

Under the spaced listening condition (E2), participants showed a significant increase in the collocation recognition section of the test from the pretest to the immediate posttest (d=1.392) and from the pretest to the delayed posttest (d=1.3404), with large effect sizes. They recognized 4.471 more collocations on average on the immediate posttest and an additional 4.059 collocations on the delayed posttest compared to the pretest. Further details on the pairwise comparison scores between pretest, immediate posttest, and delayed posttest are shown in Table 2.7.

Table 2.7. Pairwise comparison for different sections of the tests

Time of testing (*i*)	Time of testing(*j*)	Difference between means (*j*-*i*)	*SD* error	*p*	Cohen's *d*	95% confidence interval for the difference	
						Upper bound	Lower bound
A Control Mass Listening (C1)							
1	2	0.176	0.516	0.737	0.078	1.271	-0.918
	3	0.000	0.284	1.000	0.000	-0.603	0.603
B Control Mass Listening (C1)							
1	2	-0.412	0.310	0.203	0.293	0.245	-1.069
	3	-0.235	0.304	0.450	0.165	-0.408	0.879
C Control Mass Listening (C1)							
1	2	0.235	0.136	0.104	0.291	0.524	-0.054
	3	0.176	0.176	0.332	0.239	0.551	-0.198
A Experimental Mass Listening (E1)							
1	2	6.00*	0.801	<0.001	2.247	7.677	4.323
	3	5.00*	0.798	<0.001	1.917	6.670	3.330
B Experimental Mass Listening (E1)							
1	2	3.750*	0.830	<0.001	0.821	5.487	2.013
	3	2.950*	0.825	0.002	0.678	4.678	1.222
C Experimental Mass Listening (E1)							
1	2	2.350*	0.379	<0.001	1.942	3.143	1.557
	3	2.000*	0.363	<0.001	1.827	2.759	1.241
A Control Spaced Listening (C2)							
1	2	-0.385	0.350	0.293	0.310	0.377	-1.147
	3	-0.538	0.332	0.131	0.418	0.186	-1.263
B Control Spaced Listening (C2)							
1	2	0.077	0.211	0.721	0.045	0.536	-0.382
	3	-0.154	0.317	0.636	0.121	0.537	-0.845
C Control Spaced Listening (C2)							
1	2	0.077	0.178	0.673	0.122	0.464	-0.310
	3	0.077	0.265	0.776	0.109	0.653	-0.500
A Experimental Spaced Listening (E2)							
1	2	4.471*	0.654	<0.001	1.017	5.856	3.085
	3	4.059*	0.609	<0.001	1.660	5.350	2.767

B Experimental Spaced Listening (E2)							
1	2	1.941*	0.585	0.004	1.023	3.180	0.702
	3	1.647*	0.485	0.004	0.880	2.674	0.620
C Experimental Spaced Listening (E2)							
1	2	2.412*	0.272	<0.001	1.392	2.988	1.835
	3	2.294*	0.268	<0.001	1.340	2.862	1.726

Note. A= Spoken-form recognition, B= Form-meaning connection, and C= Collocation recognition.

2.7.2 Learning Gains from Massed VS Spaced Listening

In response to the second research question, investigating the effects of input distribution (massed versus spaced) on vocabulary learning, Multivariate analysis of covariance (MANCOVA) with pretest scores as the covariate was used.

2.7.2.1 Immediate Posttest

The results on the immediate posttest showed an overall statistically significant difference, F (9.00, 180.00) = 7.590, $p < 0.001$, Pillais' Trace = 0.825, partial $\eta^2 = 0.275$. The analysis exploring the difference between groups for each part of the test revealed significant differences across all three parts. Spoken-form recognition (F (3, 63) = 40.717; $p < 0.001$; partial $\eta^2 = 0.671$) and collocation recognition (F (3, 63) = 22.940; $p < 0.001$; partial $\eta^2 = 0.534$) showed extremely large effect sizes. The form-meaning connection part of the test also showed a large effect (F (3, 63) = 10.152; $p < 0.001$; partial $\eta^2 = 0.337$), but not as large as the other two parts of the test.

Post hoc comparison using the Bonferroni test was conducted. Based on the previous section of this study, no learning occurred as a result of the control group. Therefore, this section focuses on the differences between the two experimental groups. For the spoken-form recognition, E1 had significantly higher actual scores than E2 ($p = 0.001$, $d = 0.93$), which indicated that participants in the massed listening condition were better at recognizing the spoken-form of single words than those in the spaced listening condition, with a large effect size.

For form-meaning connection, E1 had significantly higher actual scores than E2 (p = 0.007, d = 1.32), indicating that participants in the massed listening condition were better at recognizing the form-meaning connection of single words than participants in the spaced listening condition, with a large effect size.

Finally, on the collocation recognition aspect of word knowledge. No significant differences were found between the experimental groups (E1 vs E2), indicating that participants did not perform any better under either massed or spaced listening conditions as far as the collocation recognition aspect of word knowledge was concerned.

2.7.2.2 Delayed Posttest

The results indicated an overall statistically significant difference, F (9.00, 180.00) = 7.505, p < 0.001, Pillais' Trace = 0.819, partial η^2 = 0.273 on the delayed posttest. The analysis exploring the difference between groups for each part of the test revealed significant differences across all three parts. Spoken-form recognition (F (3, 63) = 42.331; p < 0.001; partial η^2 = 0.679) and collocation recognition (F (3, 63) = 21.262; p < 0.001; partial η^2 = 0.515) showed extremely large effect sizes and form-meaning connection (F (3, 63) = 9.30; p < 0.001; partial η^2 = 0.317) showed a very large effect size.

The Bonferroni post hoc comparison between groups (E1and E2) for each aspect of word knowledge indicated that for the spoken-form recognition, E1 had significantly higher actual scores than E2 (p = 0.44, d = 0.739), indicating that participants in the mass listening condition continued to recognize the spoken-form of single words better than those in the spaced listening condition with a medium effect size on the two weeks delayed posttest.

For the other two aspects of word knowledge, there were no statistically significant differences between the experimental groups on the two weeks delayed posttest. Indicating that

participants did not perform any better under mass or spaced listening conditions when the form-meaning connection and collocation recognition aspect of word knowledge were concerned on the two weeks delayed posttest.

2.8 Discussion

2.8.1 Does listening to songs contribute to L2 incidental vocabulary learning?

In response to the first research question investigating the effects of repeated listening to a song on vocabulary learning, the results showed that repeated listening (i.e., Five times) could contribute to vocabulary learning. Participants showed an overall gain of 26.89% (12.1 items) from massed listening and 19.60% (8.82 items) from spaced listening conditions on the immediate posttest. On the delayed posttest, the participants retained the knowledge of 22.22% (9.41 items) in the massed listening condition and 17.77% (8 items) in the spaced listening condition after two weeks. In contrast, the control groups of both listening conditions showed no vocabulary gains.

Comparing these findings to those of Pavia et al. (2019), which found an overall gain of 8.67% (1.64 items) from the massed listening condition, the participants learned 18.22% (10.46 items) more in this study under the same condition. This is not surprising; the participants in this study had a considerably larger vocabulary size than those in Pavia et al. (2019). Research on incidental vocabulary learning from various sources of input suggests that there is a strong correlation between prior vocabulary knowledge and learning gains (Majuddin et al., 2021; Montero Perez et al., 2014; Penno et al., 2002; Peters & Webb, 2018; Webb & Chang, 2015a). Furthermore, the learning gains found in this study bear a closer resemblance to those found in other studies of incidental vocabulary learning. For example, listening to short texts in van Zeeland and Schmitt (2013b) led to a 29.2% vocabulary gain. This suggests that listening to

songs has the potential to be as effective as other sources of language input for incidental vocabulary learning, at least if the input is relatively well matched to the learners' vocabulary knowledge and, ultimately, their overall proficiency level.

2.8.1.2 Does listening to songs contribute to the learning of different aspects of word knowledge?

To evaluate the impact of massed and spaced listening to songs on different aspects of word knowledge, the knowledge of spoken-form recognition, form-meaning connection, and collocation recognition were measured separately. In the immediate posttest, massed listening led to a 31.57% improvement (6 items) on the spoken-form recognition, 19.73% (3.75items) on the form-meaning connection, and 33.57% (2.35 items) on the collocation recognition. For the spaced listening condition, the gains were 23.52% (4.47 items) on the spoken-form recognition, 10.21% (1.94 items) on the form-meaning connection, and 34.42% (2.41items) on the collocation recognition. According to the delayed posttest, massed listening condition led to 26.32% (5 items) on the spoken-form recognition, 15.53% (2.95 items) on the form-meaning connection, and 28.57% (2 items) on the collocation recognition. For the spaced listening condition the gains amounted to 21.37% (4.06 items) on the spoken-form recognition, 8.63% (1.64 items) on the form-meaning connection, and 32.71(2.29 items) on the collocation recognition.

As expected, the participants showed higher vocabulary learning gains on individual aspects of the word knowledge in comparison to those found in Pavia et al. (2019). The previous study found "10.97% (.76 words)" (Pavia et al., 2019, p.762) for the collocation recognition and no statistically significant gains for the other two aspects of word knowledge. However, their descriptive statistics showed learning gains of 11.52 % (2.19 words) on spoken-form recognition

and 2.47 % (0.47 words) on the form-meaning connection aspect of word knowledge. This means that the current study's massed listening condition, which was the replication of the previous study, led to additional learning gains of 20.05% (3.81 items) on the spoken-form recognition, 17.26% (3.28 items) on the form-meaning connection and 11.63% (0.83 items) on the collocation recognition. Considering that the difference between participants' proficiency levels was the only difference between the two studies for the massed listening condition, the higher vocabulary learning gains across different aspects of word knowledge can be attributed to better prior vocabulary knowledge. This further supports the notion that higher proficiency can lead to higher incidental vocabulary learning (Majuddin et al., 2021). Moreover, the learning gains from this study are much closer to other studies investigating incidental vocabulary learning with higher lexical coverage. For example, Peters and Webb (2018) found vocabulary learning gains of 13.95% (3.95 items) for the form-meaning recognition aspect of word knowledge from watching a TV program; furthermore, Webb and Chang (2022) found 16% (2.58 items) learning gains on the collocation recognition test from listening to an audio version of a graded reader.

2.8.2 Does the distribution of listening affect vocabulary learning gains from songs?

In response to the second research question, concerning the effects of spaced and mass listening distribution on incidental vocabulary learning, the results suggest that massed repeated listening led to greater learning gains for spoken-form recognition and form-meaning connection on the immediate posttest and higher retention of spoken-form recognition on the two-week delayed posttest. Moreover, the descriptive statistics across all parts of the tests except the collocation recognition part show higher gains for the mass listening condition than the spaced listening condition. These findings may seem contradictory to some previous research (Kim &

Webb, 2022), but they are in line with more recent findings, such as those of Serrano and Huang (2023). However, as mentioned earlier, the initial scores on the pretests differed between the mass and spaced listening groups. Using the pretest scores as covariates allowed for statistical analysis of variance between the two groups, but this does not rule out the possibility that learners in the massed learning group benefited more from the listening activity thanks to better prior knowledge (Webb & Nation, 2017). On the other hand, the participants in the spaced listening conditions retained a greater proportion of their learning gains from the immediate posttest to the delayed posttest. For example, on the spoken-form recognition test, the massed listening group scored 5.26% (1 item) lower on the delayed posttest than the immediate posttest, while the spaced listening group only regressed 2.16% (0.41 item) from the immediate posttest to the delayed posttest. The same trend is observed for the form-meaning connection test; the scores of the participants in the massed listening condition showed a regression of 4.21% (0.8 items), and the spaced listening condition showed a regression of 1.57% (0.3 items) between immediate posttest and the delayed posttest. It is important to note that although participants in the spaced listening condition showed less attrition between the immediate posttest and delayed posttest, they did not necessarily have better outcomes compared to those who were in the mass listening condition. However, the rate of attrition between spaced and mass learning conditions is consistent with previous studies (Boers, 2021; Kornell, 2009; Küpper-Tetzel et al., 2014; Nakata & Suzuki, 2019). A meta-analysis by Kim and Webb (2022) suggested that longer spacing results in more durable learning. These results are also supported by Cepeda et al.'s (2006) review of distributed practice across 317 experiments, where they found that expanding interstudy interval (ISI) produced higher retention over a longer period.

Together, these findings suggest that both massed and spaced repeated listening to songs can foster incidental vocabulary learning, with the massed condition leading to better immediate benefits, but the spaced condition possibly reduces the amount of attrition over time.

2.9 Conclusion and Direction for Future Research

This study investigated whether incidental vocabulary learning would occur through listening to songs by replicating Pavia et al. (2019)'s most effective listening condition. Of further interest was whether the distribution of practice would enhance learning gains across the three target aspects of word knowledge; spoken-form recognition, form-meaning connection, and collocation recognition. The results provided affirmative evidence that repeated listening to a song can result in incidental vocabulary learning across different aspects of word knowledge. While the effects of proficiency and prior vocabulary knowledge were not the main research objectives in this study, comparing this study's outcomes to those of Pavia et al. (2019) suggests that higher-proficiency individuals may exhibit higher learning than lower-proficiency individuals. However, more research with participants from a wider range of proficiency is needed to fully understand the effects of overall proficiency on spaced and mass listening interventions when using songs for incidental vocabulary learning. Finally, time intervals between exposures showed similar learning gains for mass and spaced learning intervals in the long term. However, due to the rate of attrition in the mass learning sessions, one may speculate that if learning occurs through spaced learning sessions it may be more durable.

This study provided further empirical evidence for the potential of songs for language acquisition. However, several limitations and possible solutions for future research need to be acknowledged. First, there were a different number of participants within each group. Initially, 172 participants were recruited, which would have resulted in 43 participants per group;

however, because of the COVID-19 pandemic and new school regulations, the study was done with smaller groups, resulting in an uneven number of participants per group. Nevertheless, the smallest group in this study is still larger than some of the other studies investigating song use for vocabulary learning (e.g., Mannarelli & Serrano, 2022; Medina, 1993). A second limitation is that his study examined only one operationalization of spacing: Participants in the spaced listening condition were exposed to the target song every 4.5 days. Research looking for optimal spacing schedules has yielded mixed results. Some suggest that longer spacing results in more durable vocabulary learning gains (Cepeda et al., 2006). However, the study by Suzuki (2017) found that 3.3-day intervals led to better retention than 7-day intervals in second language grammar learning. Therefore, further research is needed to find a more precise optimal distribution of practice for language acquisition through listening to songs. Third, this study only explored learning three aspects of word knowledge from massed and spaced listening to songs; Nation (2001, 2022) suggests that nine different aspects of word knowledge can develop under different conditions. The current study suggested that it is possible to learn the form-meaning connection through songs, a task that was previously considered challenging (Medina, 1993; Pavia et al., 2019). However, the format of the tests in this study focused on knowledge of receptive recognition alone, and it is unclear what other factors may affect the processing of word knowledge from songs that would lead to the development of other aspects, including the productive use of form-meaning connection of words. Research shows that deep processing of the meaning of lexical items that would allow for receptive and productive use likely requires encountering and using words in varied contexts with contextual clues (Teng, 2019; Webb, 2008). This brings about the fourth limitation; this study used only one song to measure vocabulary learning. Pavia et al. (2019) found different learning gains between the two songs

they used for their study; thus, further research is needed to understand better how different song characteristics may foster vocabulary learning. One way to explore this is by examining the incremental learning of the meaning of words through diverse use in different songs that offer contextual clues. Fifth, the number of target collocations (n= 7) was considerably smaller than the number of single words selected as targets (*n*= 19). Therefore, it would be beneficial to investigate incidental learning of collocations from a song that includes a higher number of collocations. This will provide results that are more comparable with other incidental vocabulary learning studies that have focused on multiword items (e.g., Majuddin et al., 2021).

In conclusion, songs remain one of the most under-researched sources of language input despite being popular with both language learners and teachers (Arnold & Herric, 2017; Peters, 2020). Therefore, further empirical studies exploring how language users utilize songs for learning, such as using songs and their lyrics in the classroom, would provide ecological validity for research with songs and provide further pedagogical implications for students and language teachers globally.

2.11 References

Arnold, J. L., & Herrick, E. (2017). *New ways in teaching with music.* TESOL International Association. Alexandria, Virginia, USA: TESOL press.

Bahrami, Z. N., Izadpanah, S., & Bijani, H. (2019). The impact of musical mnemonic on vocabulary recalling of Iranian young learners. *International Journal of Instruction, 12*(1), 977–994. https://doi.org/10.29333/iji.2019.12163a

Barcroft, J. (2002). Semantic and structural elaboration in L2 lexical acquisition. *Language Learning, 52*(2), 323–363. https://doi.org/10.1111/0023-8333.00186

Barcroft, J. (2013). Input-based incremental vocabulary instruction for the L2 classroom. In J. W. Schwieter (Ed.), *Innovative Research and Practices in Second Language Acquisition and Bilingualism* (pp. 107–138). New York, NY: John Benjamins.

Boers, F. (2021). *Evaluating second language vocabulary and grammar instruction: a synthesis of the research on teaching words, phrases, and patterns*. Routledge, Taylor & Francis Group.

Bonett, D. G. (2012). Replication-Extension Studies. *Current Directions in Psychological Science : a Journal of the American Psychological Society, 21*(6), 409–412. https://doi.org/10.1177/0963721412459512

Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., & Pashler, H. (2009). Optimizing distributed practice: theoretical analysis and practical implications. *Experimental Psychology, 56*(4), 236–246. https://doi.org/10.1027/1618-3169.56.4.236

Chen, C., & Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics, 31*(5), 693–713. https://doi.org/10.1093/applin/amq031

Dang, T. N. Y., & Webb, S. (2014). The lexical profile of academic spoken English. *English for Specific Purposes, 33*, 66–76. https://doi.org/10.1016/j.esp.2013.08.001

DeKeyser, R. M. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed., pp. 94–112). Routledge. https://doi.org/10.4324/9780203628942

González-Fernández, B., & Schmitt, N. (2020). Word knowledge: Exploring the relationships and order of acquisition of vocabulary knowledge components. *Applied Linguistics, 41*(4), 481–505. https://doi.org/10.1093/applin/amy057

Goulden, R., Nation, P., & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics, 11*(4), 341–363.

Heatley, A., Nation, P., & Coxhead, A. (2002). Range and frequency programs. Available at: http://www.victoria.ac.nz/lals/staff/paul-nation.aspx.

Horst, M., & Meara, P. (1999). Test of a model for predicting second language lexical growth through reading. *Canadian Modern Language Review, 56*(2), 308–328. https://doi.org/10.3138/cmlr.56.2.308

Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: acquiring second language vocabulary through reading. *Reading in a Foreign Language, 11*(2), 207–223.

Hulstijn, J. H. (2003). Incidental and intentional learning. In C. Doughty & M. Long (Eds.), The Handbook of second language acquisition (pp. 349–381). Malden, MA: Blackwell Publishing.

Hulstijn, J., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *Modern Language Journal, 80*(3), 327–339. https://doi.org/10.1111/j.1540-4781.1996.tb01614.x

Kim, S., & Webb, S. (2022). The Effects of Spaced Practice on Second Language Learning: A Meta-Analysis. *Language Learning*, *72*(1), 269–319. https://doi.org/10.1111/lang.12479

Köksal, O., Yağışan, N., & Çekiç, A. (2013). The Effects of Music on Achievement, Attitude and Retention in Primary School English Lessons. *Procedia, Social and Behavioral Sciences,* 93, 1897–1900. https://doi.org/10.1016/j.sbspro.2013.10.136

Kornell, N. (2009). Optimising learning using flashcards: Spacing is more effective than cramming. *Applied Cognitive Psychology, 23*, 1297–1317. https://doi.org/10.1002/acp.1537

Küpper-Tetzel, C. E., Erdfelder, E., & Dickhäuser, O. (2014). The lag effect in secondary school classrooms: Enhancing students' memory for vocabulary. *Instructional Science, 42*(3), 373–388. https://doi.org/10.1007/s11251-013-9285-2

Lai, C., Zhu, W., & Gong, G. (2015). Understanding the quality of out-of-class English learning. *TESOL Quarterly, 49*(2), 278-308. https://doi.org/10.1002/tesq.171

Mannarelli, P., & Serrano, R. (2022). 'Thank you for the music': examining how songs can promote vocabulary learning in an EFL class. *Language Learning Journal*, *ahead-of-print*(ahead-of-print), 1–15. https://doi.org/10.1080/09571736.2022.2092198

Marsden, E. (2020). *Methodological transparency and its consequences for the quality and scope of research.* In J. McKinley, & H. Rose (Eds.), The Routledge Handbook of Research Methods in Applied Linguistics (pp. 15-28). New York, New York.

Marsden, E., Morgan-Short, K., Trofimovich, P., & Ellis, N. C. (2018). Introducing registered reports at language learning: Promoting transparency, replication, and a synthetic ethic in the language sciences. *Language Learning, 68*(2), 309–320. https://doi.org/10.1111/lang.12284

Medina, S. L. (1993). The effects of music on second language vocabulary acquisition. *National Network for Early Language Learning, 6*(3), 1-8.

Montero Perez, M., Peters, E., Clarebout, G., & Desmet, P. (2014). Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology, 18*, 118–141. http://dx.doi.org/10125/44357

Muijs, D. (2011). *Doing quantitative research in education with SPSS*. London, UK: SAGE Publishers.

Majuddin, E., Siyanova-Chanturia, A., & Boers, F. (2021). Incidental Acquisition of Multiword Expressions through Audiovisual Materials: The Role of Repetition and Typographic Enhancement. *Studies in Second Language Acquisition, 43*(5), 985–1008. https://doi.org/10.1017/S0272263121000036

Nakata, T., & Suzuki, Y. (2019). Effects of massing and spacing on the learning of semantically related and unrelated words. *Studies in Second Language Acquisition, 41*, 287–311. https://doi.org/10.1017/S0272263118000219

Nation, I. S. P. (2001). Learning Vocabulary in Another Language. *In Learning Vocabulary in Another Language*. Cambridge University Press.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review, 63*, 1–12. Retrieved from http://dx.doi.org/10.3138/cmlr.63.1.59.

Nation, I. S. P. (2017). The BNC/COCA level 6 word family lists (Version 1.0.0) [Datahttp://www.victoria.ac.nz/lals/staff/paul-nation.aspx.

Nation. (2022). *Learning vocabulary in another language*. Cambridge University Press.

Newton, J. (2020). Approaches to learning vocabulary inside the classroom. In S. Webb (Ed.), *The Routledge Handbook of Vocabulary* (pp 255-270). New York, New York. doi: 10.4324/9780429291586

Pavia, N., Webb, S., & Faez, F. (2019). Incidental vocabulary learning through listening to songs. *Studies in Second Language Acquisition*, *41*(4), 745-768. doi:10.1017/S0272263119000020

Pellicer-Sánchez, A. (2016). Incidental L2 vocabulary acquisition from and while reading: An eye-tracking study. *Studies in Second Language Acquisition, 38*(1), 97–130. https://doi.org/10.1017/S0272263115000224

Pellicer-Sánchez, A. (2017). Learning L2 collocations incidentally from reading. *Language Teaching Research, 21*(3), 381–402. https://doi.org/10.1177/1362168815618428

Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do Things Fall Apart? *Reading in a Foreign Language, 22*(1), 31–55. https://doi.org/10125/66652

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. https://doi.org/10.1037/0022-0663.94.1.23

Peters, E. (2020) The effect of out-of-class exposure on vocabulary knowledge. In S. Webb (Ed.) *Approaches to learning, testing, and researching L2 vocabulary*. John Benjamins Publishing Company (pp.143-167). https:doi.org/10.107/bct.109.itl.00010.pet

Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, *40*(3), 551–577. https://doi.org/10.1017/S0272263117000407

Pigada, M., & Schmitt, N. (2006). Vocabulary Acquisition from Extensive Reading: A Case Study. *Reading in a Foreign Language*, *18*(1), 1–28.

Porte, G., & McManus, K. (2019). *Doing replication research in applied linguistics*. Routledge. New York, NY: Taylor & Francis.

Reynolds, B. L. (2015). The effects of word form variation and frequency on second language incidental vocabulary acquisition through reading. *Applied Linguistics Review, 6*, 467–497. https://doi.org/10.1515/applirev-2015-0021

Rodgers, M.P.H., & Webb, S. (2020). Incidental vocabulary learning through watching television. *ITL - International Journal of Applied Linguistics*, *171*(2), 191-220. https://doi.org/10.1075/itl.18034.rod

Rogers, J. (2015). Learning second language syntax under massed and distributed conditions. *TESOL Quarterly, 49,* 857–866. https://doi.org/10.1002/tesq.252

Roohani, A., & Akbarpour, S. (2016). The effect of using songs on young Iranian EFL learners' vocabulary learning: A case of gender. *English Language Teaching, 3*(3), 59-82.

Rott, S. (1999). The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition and retention through reading. *Studies in Second Language Acquisition, 21*, 589–619. https://doi.org/10.1017/S0272263199004039

Saragi, T., Nation, P., & Meister, F. (1978). Vocabulary learning and reading. System, 6, 72–78. Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual.* Basingstoke, UK: Palgrave Macmillan.

Serrano, R., & Huang, H-Y. (2018). Learning vocabulary through assisted repeated reading: How much time should there be between repetitions of the same text? *TESOL Quarterly, 52*, 971–994. https://doi.org/10.1002/tesq.445

Serrano, R., & Huang, H. (2023). Time distribution and intentional vocabulary learning through repeated reading: a partial replication and extension. Language Awareness, 32(1), 1–18. https://doi.org/10.1080/09658416.2021.1894162

Suzuki, Y. (2017). The optimal distribution of practice for acquisition of L2 morphology: A conceptual replication and Extension. *Language Learning, 67*(3), 512-545. doi:10.1111/lang.12236

Tegge, F. (2015). *Investigating song-based language teaching and its effect on lexical learning* (Unpublished doctoral dissertation). Victoria University of Wellington, New Zealand. doi: 10.13140/RG.2.1.3159.8560.

Teng, F. (2018). Incidental vocabulary acquisition from reading-only and reading-while listening: A multi-dimensional approach. *Innovation in Language Learning and Teaching,12*, 274–288. https://doi.org/10.1080/17501229.2016.1203328

Teng, F. (2019). The effects of context and word exposure frequency on incidental vocabulary acquisition and retention through reading. Language Learning Journal, 47(2), 145–158. https://doi.org/10.1080/09571736.2016.1244217

Tilwani, S. A., Amini MosaAbadi, F., Shafiee, S., & Azizi, Z. (2022). Effects of Songs on Implicit Vocabulary Learning: Spoken-Form Recognition, Form-Meaning Connection, and Collocation Recognition of Iranian English as a Foreign Language Learners. *Frontiers in Education*, *7*: (797344). https://doi.org/10.3389/feduc.2022.797344

Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language Learning*, 69(3), 559-599. https://doi.org/10.1111/lang.12343

Van Zeeland, H., & Schmitt, N. (2013a). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System*, *41*, 609-624. doi: http://dx.doi.org/10.1016/j.system.2013.07.012

Van Zeeland, H.,&Schmitt, N. (2013b). Lexical coverage and L1 and L2 listening comprehension: The same or different from reading comprehension. *Applied Linguistics, 34,* 457–479. doi: 10.1093/applin/ams074

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language, 15*, 130–163. https://doi.org/10.1177/003368828501600214

Warner. R. M. (2013). Applied statistics : from bivariate through multivariate techniques (2nd ed.). *SAGE Publications.*

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics, 28*(1), 46–65. https://doi.org/10.1093/applin/aml048

Webb, S. (2008). The Effects of Context on Incidental Vocabulary Learning. Reading in a Foreign Language, 20(2), 232–245.

Webb, S. (2020). *Incidental vocabulary learning*. In S. Webb (Ed.), *The Routledge Handbook of Vocabulary* (pp 255-270). New York, New York. doi: 10.4324/9780429291586

Webb, S. A. , & Chang, A. C-S, (2012a). Second language vocabulary growth. *RELC Journal, 43*, 113–126. doi: 10.1177/0033688212439367.

Webb, S. A., & Chang, A. C-S. (2012b). Vocabulary learning through assisted and unassisted repeated reading. *Canadian Modern Language Review, 68*, 267–290. doi: 10.3138/cmlr.1204.1.

Webb, S., & Chang, A. C. S. (2015a). How does prior word knowledge affect vocabulary learning progress in an extensive reading program*? Studies in Second Language Acquisition, 37,* 651–675.

Webb, S., & Chang, A. C-S. (2015b). Second language vocabulary learning through extensive reading: How does frequency and distribution of occurrence affect learning*? Language Teaching Research, 19,* 667–686.

Webb, S., & Chang, A. C-S. (2020). How does mode of input affect the incidental learning of collocations? *Studies in Second Language Acquisition (2020),* 1–22. doi:10.1017/S0272263120000297

Webb, S., & Chang, A. C.-S. (2022). How does mode of input affect the incidental learning of collocations? *Studies in Second Language Acquisition, 44*(1), 35–56. https://doi.org/10.1017/S0272263120000297

Webb, S., & Nation, P. (2017). *How vocabulary is learned.* Oxford, UK: Oxford University Press.

Webb, S., Yanagisawa, A., & Uchihara, T. (2020). How Effective Are Intentional Vocabulary-Learning Activities? A Meta-Analysis. *The Modern Language Journal (Boulder, Colo.), 104*(4), 715–738. https://doi.org/10.1111/modl.12671

Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *Canadian Modern Language Review, 57*(4), 541–572. https://doi.org/10.3138/cmlr.57.4.541

Chapter 3: Study 2

Receptive and Productive Learning of Formulaic Sequences from Spaced Listening to Songs: The Role of Different Modes of Input

3.1 Abstract

In this study, 98 learners of English as a foreign language at a secondary school in Thailand were randomly assigned to a control, or one of four experimental groups to investigate the effects of multimodal song use on incidental vocabulary learning. The control group completed a pretest, immediate posttest, and two-week delayed posttest of receptive (form recognition) and productive (form recall) vocabulary tests without exposure to target vocabulary items. Each experimental group was exposed to 16 target formulaic sequences (FS) in the same song in one of the four following modes of input: (I) listen only (L), (II) listen while reading the lyrics (LL), (III) listen and sing (LS), (IV) listen and sing while reading the lyrics (LSL). The results showed statistically significant learning gains for all experimental groups but no significant learning gains for the control group on the immediate and the two weeks delayed posttest on both the productive and receptive vocabulary tests. Overall, the experimental groups outperformed the control group, while the LSL group outperformed all the other groups. Furthermore, the LL group had the second-highest learning gains on the productive tests, followed by LS and L groups. However, the difference between the LS and L was not statistically significant. On the receptive tests, the only significant difference between the experimental groups was LSL outperforming LL.

Keywords: Incidental vocabulary learning; song; receptive knowledge; productive knowledge; formulaic sequences; spaced listening; modes of input

3.2 Introduction

Many empirical studies across disciplines in the last few decades have identified high percentages of pre-constructed phrases or multi-word sequences in language discourse (Sinclair, 1991; Nelson, 2018; Howarth, 1998; Wray, 2000). For second/foreign language learners to reach high proficiency levels in any target language, they would need to learn the varied pre-constructed phrases found in that language (Boers et al., 2006; Siyanova-Chanturia & Pellicer-Sánchez, 2019). However, research suggests that even advanced English as a second/foreign language ESL/EFL students struggle with collocations, a type of formulaic language, and continue to make mistakes when using multi-word sequences (Laufer &Waldman, 2011). This

suggests that more research is needed to find more efficient formulaic language learning methods.

Available studies investigating second language acquisition have come to the agreement that deliberate teaching of collocations and teaching through meaning-focused activities have the potential to foster the acquisition of collocations in second/foreign language learners (Boers & Webb, 2018; Laufer, 2003; Nation, 2013, Webb & Nation, 2017). In recent years, research has shown that incidental learning through different input sources can also contribute to collocation acquisition. Webb and Chang (2022), investigating the learning of 17 different collocations from different modes of input, found positive learning gains from reading, listening, and reading while listening to a graded reader. Their results concluded that reading while listening was the most effective mode of learning, followed by smaller learning gains through listening and, finally, reading. These findings highlighted the critical role of listening for collocation acquisitions. In comparison to studies investigating incidental learning of single words, where listening contributed the least amount of learning gains (Brown, Waring, & Donkaewbua, 2008; Webb & Chang, 2012; Webb, Newton, & Chang, 2013), listening has been shown to play an essential role for collocation learning (Webb & Chang, 2022).This may be due to the prosodic aspect of spoken language, which makes word pairings more salient for listeners (Lin, 2012).

Similar patterns of learning were found by Pavia et al. (2019), investigating incidental vocabulary learning of single-words and collocations through listening to songs. In this study, participants showed higher learning levels for collocations (10.97%) compared to single-word items (6.53%). Their findings are not surprising as their learning conditions were entirely dependent on the listening mode of input. Teachers commonly use songs to teach single-word and multi-word sequences (Arnold & Herrick, 2017). However, Pavia et al. (2019) is the first

empirical study investigating incidental collocation acquisition from songs. Thus, more research is needed to examine further how songs could potentially foster the learning of formulaic language.

The present study examined how different modes of input (Listening, listening while reading the lyrics, listening while singing, and finally listening and singing while reading the lyrics) influence the receptive and productive learning of formulaic language from songs over five sessions. Furthermore, previous research suggests that distributed practice positively affects vocabulary learning and leads to more durable learning over time compared to mass exposure (Kim & Webb, 2022); therefore, listening sessions for the experimental groups were spaced by 4.5 days.

3.3 Literature review

3.3.1 Receptive language learning through songs

Receptive language knowledge allows individuals to understand a written text or listening task (Webb, 2008). Gary and Gary (1981) described five benefits of focusing on receptive listening teaching/learning at the beginning stages of language learning. First, receptive knowledge develops faster than productive knowledge, which will allow learners to experience and enjoy more of the target language in a shorter period. Second, this speed of development can be very motivating. Third, from a cognitive point of view, focusing on one skill (e.g., listening) at a time will not overwhelm students. Fourth, productive use of language can be very stressful; thus, focusing on receptive use can have psychological benefits and reduce foreign language class anxiety (FLCA), leading to fewer unwanted student behaviours (Horwitz et al., 1986). Fifth, receptive learning through listening can be done independently; thus, students can engage

in receptive language learning activities, including listening to songs outside the classroom and on their own.

Listening to songs in their most natural form is a receptive task without any deliberate instructions. Thus, it is essential to explore the extent to which learning can occur from merely listening to songs. A few studies have explored this. However, their experimental designs were limited to a small range of songs; two pop songs (Pavia et al., 2019) or children's storybook songs (Medina, 1993). Besides, the participants of these studies were limited to children under the age of 14. Thus, several variables still need to be considered and further researched.

Furthermore, using multiple receptive skills (reading and listening) has been shown to increase learning gains from other input sources. For example, Webb and Chang (2022), investigating vocabulary learning gains from reading, listening, and reading while listening, found that reading while listening led to higher levels of learning on both immediate and delayed posttests. A study of learning vocabulary through audiovisual input (watching TV programs) with and without on-screen texts (captions) demonstrated higher levels of learning in participants who watched the program with captions (Peters, Heynen, & Puimege, 2016). Teachers and students have used songs and their unsimplified lyrics as learning materials for decades (Arnold & Herrick, 2017; Tegge, 2015). Resources such as the Wellington Corpus of Popular Songs (WOP), Wellington Corpus of Popular Songs in English Teaching (WOPET) (Tegge, 2015) and the ability to add a caption to any music video on websites such as www.Youtube.com have made use of songs with their lyrics very accessible and widely used by learners and educators around the world. Thus, it would be useful to investigate to what extent listening to songs while reading the lyrics can contribute to receptive knowledge of new words.

3.3.2 Productive language learning through songs

Productive language knowledge allows individuals to speak and write (Webb, 2008). Receptive knowledge can contribute to productive knowledge and provide insight into their productive knowledge (Webb, 2008; Nation & Webb, 2011). However, several studies investigating the receptive-productive knowledge relationship focusing on vocabulary size have indicated that receptive vocabulary size exceeds productive vocabulary size (Morgan & Oberdeck, 1930; Laufer, 1998; Webb, 2008). This suggests that receptive knowledge does not automatically transfer to productive knowledge, and as Swain (1985; 2005) argues, productive language learning is essential for productive language development.

Singing and sing-along activities are commonly practiced around the world and are characterized as fun and enjoyable activities. Since the invention of the "sing-along system" or "Karaoke" (Zhou & Tarocco, 2007, p.149) by Del Rosario in 1978, karaoke has become a global phenomenon. In 2018 more than 86,000 Karaoke machines were sold in the United States alone, equating to a wholesale price of 13.22 million US dollars (Lock, 2019). Karaoke machines are not the only devices that allow for sing-along activities. Free video streaming websites, including www.Youtube.com, have channels dedicated to sing-along activities with millions of songs to choose from. Many ESL/EFL teachers have used these resources as a medium through which students can learn and practice language productively through singing (Arnold & Herrick, 2017). For decades, teachers and educators have suggested step-by-step instructions on using songs and singing in the language classroom (Guglielmino, 1986; Arnold & Herrick, 2017). However, the extent to which sing-along activities affect second/foreign language learning outcomes is still unclear.

The productive use of language in sing-along activities can be administered and measured in different ways. First, the learners can be exposed to a song and be asked to sing along without any other aiding materials. This condition mimics what individuals do naturally if they hear their favourite song on the radio or attend a musical concert. Second, learners can be exposed to the song's melody and lyrics like a Karaoke machine. Third, in the attempt to integrate the first two conditions, learners can be exposed to the song with the singer singing the words and singing along while reading the lyrics, similar to what language teachers have reported doing in language classrooms (Arnold & Herric, 2017).

These teaching/learning conditions make use of receptive and productive language use. Thus, both receptive and productive tests can be used to evaluate different aspects of language knowledge, including the recognition and production of formulaic sequences used to measure language learning in this study.

3.3.3 Focus on Formulaic Sequences (FS)

An abundance of research has investigated Formulaic language. However, over 50 different terms have been used to refer to the same phenomenon (Wray, 2002), including; "fixed expressions, formulaic language, conventionalized forms, lexicalized phrases, prefabricated routines" (Siyanova-Chanturia & Pellicer-Śanchez, 2019. p.2) to name a few. The present paper will use the inclusive term "Formulaic sequences (FS)" (Wray, 2001) to define the underlying phenomenon that "certain words have a strong relationship with each other in creating meaning" (Siyanova-Chanturia & Pellicer-Śanchez, 2019. p.3). This umbrella term includes "lexical bundles and n-grams (in the meantime, I don't know), multi-word verbs (catch up), collocations (spread the news), irreversible binomials (bride and groom), idioms (it came straight from the

horse's mouth), and proverbs (A rolling stone gathers no moss)" (Siyanova-Chanturia & Pellicer-Śanchez, 2019. p.38).

Formulaic sequences make up a large proportion of the English language. Erman and Warren (2000) indicated that approximately 59% of oral language and 52% of written language is composed of prefabricated language. Hill (2001) reported approximately 70% collocation across different language discourses. These findings suggest that dominant/L1 language users possess large repertoires of pre-constructed phrases they draw from in daily language use. For ESL/EFL learners, learning FS remains a struggle. Research shows that the lack of knowledge and reliance on multi-word sequences contributes to the lack of proficiency and fluency in the target language in second/foreign language learners (McCauley & Christiansen, 2017). Thus, knowledge of FS can be useful at every stage of proficiency. For beginners memorizing simple, useful FS allows them to quickly gain fluency to communicate in the target language (Newton & Nation, 2020). For more advanced learners, knowledge of FS will assist with the online processing of information, leading to more effortless understanding and use of the target language (Siyanova-Chanturia & Pellicer-Śanchez, 2019). Taken together, further research on learning formulaic sequences for L2/FL learners is warranted.

3.2.4 Learning Formulaic Sequences from Songs

Lin (2012) argues that "prosody underlies the mechanism by which we learn and remember formulaic language" (p.343). Children's learning of their first language is highly influenced by what they hear from their surroundings. They imitate the prosody of formulaic language from adults and master it before learning other aspects about these formulaic chunks, such as orthography or morphology. González Fernández and Schmitt (2015) found that everyday engagement, such as out-of-school reading, watching television and movies, showed a

stronger correlation between knowledge of collocation than the number of years learning the target language. Thus, language learners, especially in a foreign context where their primary sources of language input are text-dominated, will not have many opportunities to be exposed to authentic oral language. Hence, they will not experience prosody-driven learning and will thus struggle with learning FS (Lin, 2012).

Empirical studies investigating incidental learning of collocations demonstrated higher learning gains through listening than reading (Webb & Chang 2022). These findings can be explained by the critical role of the prosody of speech encountered in listening tasks. Therefore, more authentic listening materials such as songs can be a valuable source of input that can promote the learning of FS in EFL contexts. While very little research demonstrates FS learning from songs (Tegge, 2015; Pavia et al., 2019), foreign/second language teachers have used the prosody-learning mechanism from songs in language teaching. In an extensive scale survey of 568 language teachers, Tegge (2018) found that 56% of the respondents use songs for teaching pronunciation and prosody in the target language. Therefore, more empirical research is needed to understand how songs can be utilized to learn formulaic sequences.

3.4 Current Study

Research on out-of-school exposure to foreign languages continues to show that listening to English songs is the number one activity among language learners. However, these studies show an inconsistent relationship between language learning and listening habits to songs (González-Fernández & Schmitt, 2015; Lindgren & Munoz, 2013; Peters, 2020). Thus, exploring how these listening habits can be utilized more efficiently to foster language learning through empirical studies is important. Evidently, the empirical study by Pavia et al. (2019) showed that repeated listening could lead to vocabulary learning from listening to songs and that learning

gains may differ from song to song. This suggests that exposure alone may not be enough for language learning to occur, and other factors that could enhance the quality of engagement should be considered (González-Fernández, 2022).

The present study was designed to explore how different receptive (i.e., Listen only, and listen while reading the lyrics) and receptive in addition to productive (i.e., Listen while singing, and listen and sing while reading the lyrics) use of a song enhances vocabulary learning. Receptive use was operationalised here as exposure to the target song through listening and reading. Production was operationalized as vocally imitating aural input (i.e., singing existing lyrics). Vocabulary learning was measured by comparing scores on a formulaic sequences productive recall test and a recognition test from the pretest to the immediate posttest and delayed posttest. Considering the strong evidence for the positive effects of spaced practice (Kim & Webb, 2022; Macis et al., 2021) and repeated exposure (Pavia et al., 2019) on learning, L1 Thai participants in this study were exposed to a single song five times over three weeks.

3.5 Research Questions

3.5.1 Measuring Overall FS Learning

Does spaced listening to a song contribute to L2 formulaic sequences learning?

3.5.1.1 Measuring Receptive and Productive Knowledge Separately

What are the receptive and productive FS learning gains from repeated exposure (5 times) to a song over five sessions?

3.5.2 Comparing Receptive and Productive Activities

Which learning condition leads to higher learning gains over five sessions?

Repeated exposure (5 times) to a song through:

1. Listen first, then listen again four times

2. Listen first, then listen while reading the lyrics four times
3. Listen first, then listen and sing along four times
4. Listen first, then listen, read the lyrics, and sing along four times

3.6 Method

3.6.1 Participants

The participants were 98 Thai-speaking students in Matthayom four and five (Grades 10 &11), with ages ranging from 15 to 18 (M = 16.31, SD = 0.66). Most of the participants had taken English as a compulsory subject in school for approximately 12 years, starting at the pre-elementary/preschool level. Attending preschool is not mandatory in Thailand; however, more than 74% of Thai students attend preschool education (Bureau of International Cooperation, 2008). The school had estimated that the participants' proficiency was at the beginner-intermediate level. Based on the results of the updated Vocabulary Levels Test (VLT) (Webb et al., 2017)(See Appendix D), participants had substantial knowledge of the form-meaning connection of the most frequent 1000-word families (M = 26.27 out of 30, SD =2.7) and were learning the most frequent 2000 (M = 15.10 out of 30, SD = 6.83) and 3000 (M = 6.46 out of 30, SD = 4.29) word families. Based on Webb and Chang (2015a), a correct response on the VLT represents knowledge of 33.3 words; thus, the participants in this study had knowledge of approximately 1,593 of the most frequent 3,000-word families.

The Thai school had randomly assigned students to one of five classes within their respective grades (i.e., Grade 10 and 11). This study randomly assigned each class to one of the five conditions (i.e., C, E1, E2, E3, & E4). The results of the Kolmogorov-Smirnov test of normality were non-significant (p = 0.065) based on the participants' VLT scores and confirmed that they were normally distributed across conditions.

3.6.2 Research Instruments

3.6.2.1 Song

The clean radio version of the song "Circles" by Post Malone (2019) was selected as the learning material for this study. Four criteria were considered for selecting this song. First, this song had to be new to the participants based on communications with the school and Thai students in the same age group as the target participants in this study. The song's novelty ensured that the learning gains from the study are attributed to the study's learning conditions (Nation & Webb, 2011). Second, the selected song had to include formulaic sequences likely unknown to the target participants. These target FS in the target song were of diverse types (i.e., collocations, lexical bundles, collocations, phrasal verbs, etc.). This reflects the diversity of FS in natural discourse. (The aim of this study was not to investigate if some types of FS stand a better chance than others to be picked up incidentally from songs.) Third, the song needed to be age-appropriate with no vulgar language, making it appropriate for educational use. Fourth, to ensure optimal learning, previous studies suggest that 95% of the vocabulary items in the learning materials need to be within the learner's previous knowledge (Van Zealand & Schmitt, 2013). Thus, the lyrics of the target song (see Appendix H) were analyzed through the updated *Range* program available on Cobb's (2020) website https://www.lextutor.ca/cgi-bin/range/texts/index.pl using the British National Corpus/Corpus of contemporary American English (BNC/COCA). The results (See Table 3.1) suggested that participants would need to know between 1,000 and 2,000 most frequent word families to reach 93.09%-98.80% lexical coverage to learn target words incidentally. Compared to the target participants' knowledge in this study, with knowledge of approximately 1593 most frequent word families, it can be speculated that they reached 95%

lexical coverage; therefore, this song was at the appropriate difficulty level for unassisted incidental vocabulary learning to occur.

Table 3.1

Lexical frequency Profile: “Circles.”

Level	Token%	Cumulative coverage %
1	93.09	93.09
2	5.71	98.80
3	0.60	99.40
4	0.00	99.40
5	0.60	100.00

3.6.2.2 Target Formulaic Sequences (FS)

Twelve dominant English speakers were given a brief definition of formulaic sequences (FS) and were asked to identify those found in this study's target song's lyrics. The responses ranged from 17 to 19. Sixteen FS (see Table 3.3) were identified by every individual and thus selected and used in this study. Mutual information (MI) scores, retrieved from the COCA (1990-2019) corpus of approximately three and above, were considered for the target collocations to ensure the probability of two words occurring together would be higher than them occurring independently (Beran, 2000). However, given that MI scores cannot be applied to high-frequency function words, they were still considered for the study based on their high corpus frequency (see Table 3.3) and recognition by lexicographers as conventionalized expressions. For example, “Every time” appears as an idiom in the Merriam-Webster dictionary (Retrieved from: https://www.merriam-webster.com/dictionary/every%20time).

It was mentioned previously that FS stands a comparatively good chance of being picked up from aural input owing to their prosodic features. For example, speakers tend to pause between FS, not within them, which can help learners identify such word strings as lexical units. It is important to note that how formulaic sequences are used in speech may differ from how they are used in songs because of considerations of rhythm and melody, and so the way FS are sung does not always correspond to their prototypical prosody in speech. However, the way the selected FS for this study were produced in the song either corresponded to their prototypical prosody in speech or enhanced the prosodic features that can help make FS salient. For example, "turn around" in "Would it really matter when they could **turn around** and tax you on something else?" from an interview was printed in the Philadelphia City Paper (2010) follows the same prosody in speech as the corresponding FS in the current study. In contrast, "run away" in "when adopted that the dog had a tendency to **run away**, she said." from an interview on Seward City News (2012) is less prosodic than the corresponding FS in the song used for the current study since the prosodic boundary markers are more salient in this particular song.

Table 3.2 Target FS

Song	Target FS	Number of encounters in the song	*fr*	MI Score
Circles	Turn Around	1	28591	4.61
	Upside Down	1	5843	7.40
	Feed the Flame	2	86	2.54
	Every Time	1	38792	-
	Seasons Change	2	104	3.73
	Let Go/ Let it Go	4	14028	5.81
	Run Away	6	7031	8.49

Running in Circles	2	233	3.50
Take the Blame	2	1154	2.73
Got a Feeling	2	605	-
From the Get-go	1	996	-
I'm going Through	1	19965	2.54
Waiting On	2	3295	-
Don't understand	1	3574	-
I said so	1	12612	-
Make up your mind	1	654	-

3.6.2.3 Dependent Measures

A vocabulary test was created to measure the learning gains of the participants throughout the study. This test was used for all groups during the pretest, immediate posttest, and two-week delayed posttest (see Appendix I).

The test included two sections. One measured receptive knowledge, while the other measured productive knowledge of FS. Previous research investigating receptive and productive vocabulary learning suggests that receptive learning conditions can lead to receptive knowledge of target words, while these conditions may not have a significant impact on productive knowledge. Meanwhile, productive learning conditions can potentially promote greater learning gains in receptive and productive knowledge (Webb, 2009). In this study, all participants completed both receptive and productive sections of the test. Thus, the assumption that participants in the productive learning conditions would outperform the participants in the receptive learning condition was explored.

In both the receptive and productive sections of the test, participants were exposed to the written and spoken format of the test, meaning they saw each question on the paper while hearing the spoken format from a recording. This way, the type of exposure stayed consistent between treatment and testing.

The first section, fill-in-the-blank, measured the productive knowledge of FS as the participants were asked to produce the written form of missing words in the provided lyrics. The sentences were taken directly from the lyrics of the target song. For scoring, the overall shape of the response was graded. Participants' responses were not marked incorrect over minor spelling mistakes. For example, "Circels" instead of "Circles" was marked as correct (Webb, 2009). The second section, matching, measured FSL recognition through a multiple-choice format test. The choices include one correct answer, two distracters, and one "I do not know" in Thai to prevent students from guessing. In accordance with the guidelines outlined by Nation and Webb (2011), one distractor was a correct answer to another question on the test. The second distractor was a word from the target song but not an answer to any of the questions. Having distractors from the learning material ensured that the students did not dismiss the choice solely due to lack of exposure. The goal was to select the correct answer based on the connection between the different words in the formulaic sequences.

3.6.3 Procedure

The five groups in this study were divided into three sections, the control group (C), the receptive learning condition groups (E1-E2), and the receptive and productive learning condition groups (E3-E4). The study took six weeks and eight sessions to complete. Table 3.3 provides an overview of the procedure timeline. During week one, the procedure described in chapter one

regarding collecting consent forms, demographic information, VLT, and the pretest, was implemented.

Between sessions two and five, the control group (C) returned to their regular Thai school curriculum. During session six (Week 4), they completed the immediate posttest. In the last session on week six, they completed the two-week delayed posttest. The control group participants were not exposed to the target songs at any point in the study.

The experimental groups (E1-E4) listened to the target song one time during week one, session two. From session three to session six, these experimental groups were exposed to one of the learning conditions once every session. The learning conditions for each experimental group are as follows; (E1) Listening only to the target song (L), (E2) Listening while reading the lyrics (LL), (E3) Listening while singing the song out loud (LS), (E4) Listening while reading and singing out loud (LSL). During session six, the participants in experimental groups E1-E4 completed the posttest immediately after their last exposure to their respective treatment. In the last session, session eight, all the participants across all experimental groups completed the two-week delayed posttest.

For the duration of the study, two approaches were taken in responding to student questions. First, if the questions could have compromised the validity of the results in any way, the participants were told that they would receive the answers at the end of our last session together. Some examples of these types of questions include clarifying the meaning of a specific word in a song or identifying the name of the song or artist. Second, if the questions asked were about what was expected of the participants and clarification of the instructions, the questions were answered right away. For example, some participants inquired whether they could work on

the questions together or utilize their electronic devices for assistance. It was clarified that all tasks were meant to be completed individually, and using electronic devices was not allowed.

Table 3.3 Procedure Timeline

	Group	Week1	Week 2		Week 3	Week 4		Week 5	Week6
		Session 1	Session 2	Session 3	Session 4	Session 5	Session 6	Session 7	Session 8
Control (C)		• Consent Form • Demographic information form • VLT • Pretest	Regular Thai school curriculum	Regular Thai school curriculum	Regular Thai school curriculum	Regular Thai school curriculum	Immediate posttest	Regular Thai school curriculum	Delayed posttest
Receptive Learning conditions	Listen Only (L)		Listened to the target song once	Listened once	Listened once	Listened once	Listened once followed by the immediate posttest		
	Listen with Lyrics (LL)		Listened to the target song once	Listened while reading the lyrics once	Listened while reading the lyrics once	Listened while reading the lyrics once	Listened while reading the lyrics once followed by the immediate posttest		
Receptive and Productive learning conditions	Listen and Sing (LS)		Listened to target song once	Listened and sang once	Listened and sang once	Listened and sang once	Listened and sang once, followed by immediate posttest		
	Listen and Sing with Lyrics (LSL)		Listened to target song once	Listened and sang while reading the lyrics once	Listened and sang while reading the lyrics once	Listened and sang while reading the lyrics once	Listened and sang while reading the lyrics once, followed by the immediate posttest		

3.6.4 Analysis

The numerical data in this study was analyzed using quantitative statistical analysis (Mujis, 2011). *IBM SPSS Statistics* version 28 data analysis generated descriptive statistics such as means, standard deviation, etc., and inferential statistics, including the multivariate analysis of variance (MANOVA), to compare the results between pretest, posttest, and delayed posttest within and between groups. Considering that the participants were randomly assigned to different groups by the Thai school, the Kolmogorov-Smirnov test of normality was also used to ensure that the participants were normally distributed between groups with respect to their proficiency level based on their scores on the VLT during session one. Mauchly's Sphericity test was used to assume homogeneity of within-group variance. If the assumption of sphericity had been violated, the Greenhouse-Geisser was used to adjust the *df* (Warner, 2013).

3.7 Results

The descriptive statistics for formulaic sequences scores on the overall FS knowledge is presented in Table 3.4. In addition, Table 3.5 illustrates the participants' scores on the receptive and productive aspect of FS knowledge separately. Repeated measures ANOVA was used to answer the first research question investigating the interaction between time and exposure to a song under each receptive and productive listening condition.

Table 3.4 Descriptive Statistics of Overall Vocabulary Test Scores: Means (and SDs)

Participant subgroups	Total FS Knowledge		
	Pretest	Immediate posttest	Delayed posttest
Control (C) (*n= 21)*	2.57 (1.57)	2.76 (1.41)	2.57 (1.47)
Listen only (L) (*n= 21)*	2.76 (1.81)	14.81 (2.82)	8.81 (2.36)
Listen while reading the lyrics (LL) (*n = 15)*	2.47 (1.92)	16.20 (4.26)	13.07 (3.97)
Listen and sing (LS) (*n= 21)*	2.38 (1.43)	14.19 (3.57)	12.29 (3.05)
Listen and sing while reading the lyrics (LSL) (*n= 20)*	3.05 (1.86)	21.50 (4.97)	20.80 (4.61)

Note. The maximum score combining both receptive and productive knowledge was 32.

Table 3.5 Descriptive Statistics of Receptive and Productive Vocabulary Test Scores: Means (and SDs)

Participant subgroups	Productive FS Knowledge			Receptive FS Knowledge		
	Pretest	Immediate posttest	Delayed posttest	Pretest	Immediate posttest	Delayed posttest
Control (C) (*n= 21)*	0.71 (0.85)	0.76 (0.83)	0.67 (0.86)	1.90 (1.09)	2.00 (1.05)	1.90 (1.09)
Listen only (L) (*n= 21)*	0.81 (0.93)	3.86 (1.24)	1.86 (1.01)	1.95 (1.02)	10.95 (1.83)	6.95 (1.75)
Listen while reading the lyrics (LL) (*n = 15)*	0.60 (0.74)	6.64 (2.35)	5.20 (1.93)	1.87 (1.41)	9.53 (2.03)	7.87 (2.20)
Listen and sing (LS) (*n= 21)*	0.52 (0.75)	4.05 (1.63)	3.67 (1.42)	1.86 (1.01)	10.14 (2.03)	8.62 (1.80)
Listen and sing while reading the lyrics (LSL) (*n= 20)*	0.70 (1.04)	9.75 (2.90)	9.35 (2.68)	2.35 (1.09)	11.75 (2.47)	11.45 (2.26)

Note. The maximum score on the productive test and the receptive test was 16.

3.7.1 Overall FS learning from spaced listening to a Song

As shown in Figure 3.1, no learning gains were made between the pretest, posttest, and delayed posttest for the control (C) group. In contrast, the experimental groups all showed learning gains between the pretest, posttest, and delayed posttest. The analysis for the control group showed that the assumption of sphericity had been met $\chi^2(2) = 1.554$, $p = 0.460$, and the within-participant main effect (time) was not statistically significant F (2, 40) = 2.443, $p = 0.10$.

The analysis for the listening-only (L) group showed that the assumption of sphericity had been met, $\chi^2(2) = 3.355$, $p = 0.187$, and the within-participant main effect (time) was statistically significant, F (2, 40) = 384.36, $p < 0.001$ with partial $\eta^2 = 0.951$ (Very large effect size). The assumption of sphericity had been violated for the listen while reading the lyrics (LL) ($\chi^2(2) = 14.633$, $p < 0.001$), listen and sing (LS) ($\chi^2(2) = 19.319$, $p < 0.001$), and listen and sing while reading the lyrics (LSL) ($\chi^2(2) = 40.026$, $p < 0.001$) groups; therefore, a Greenhouse-Geisser correction was used. The within-participant main effect (time) indicated significant results for all three groups; LL, $F(1.194, 16.711) = 218.812$, $p < 0.001$ with partial $\eta^2 = 0.940$ (Very large effect size), LS, $F(1.221, 24.416) = 338.967$, $p < 0.001$ with partial $\eta^2 = 0.944$ (Very large effect size), and LSL, $F(1.057, 20.087) = 229.969$, $p < 0.001$ with partial $\eta^2 = 0.924$ (Very large effect size).

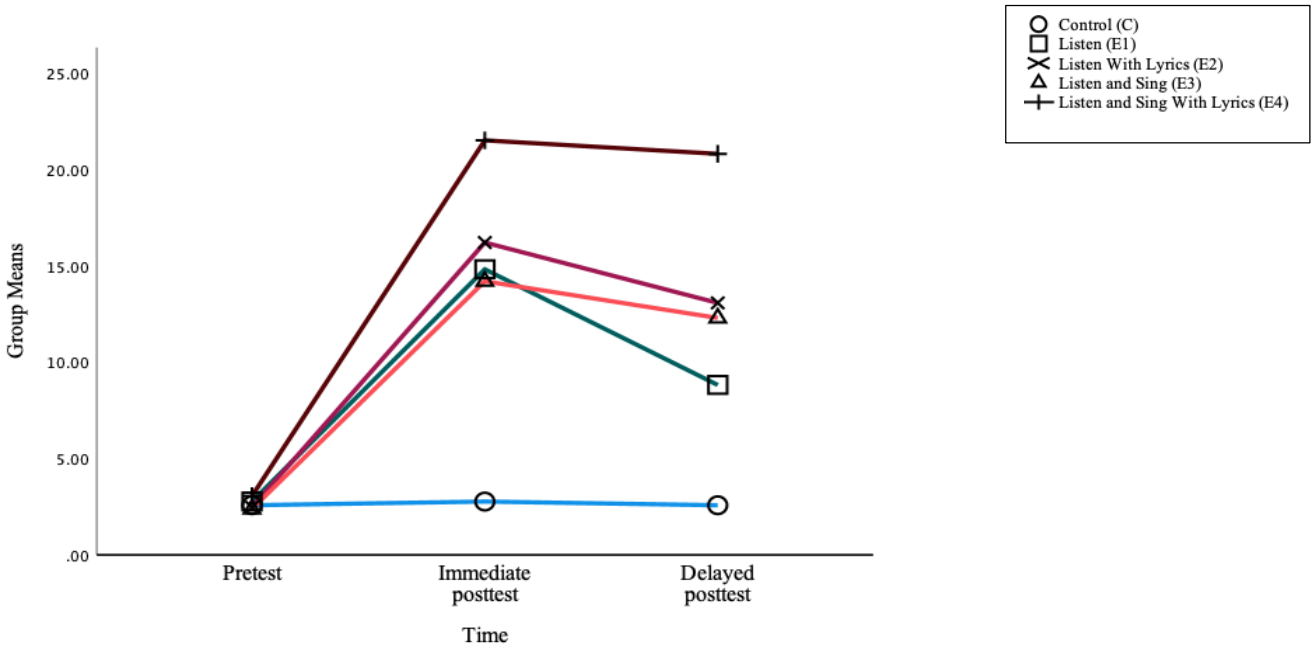


Figure 3.1 Group means for overall vocabulary learning over time.

3.7.1.1 Measuring Receptive Knowledge of FS

Figure 3.2 shows that no learning gains were made between the spoken/written form recognition pretest, posttest, and delayed posttest for the control (C) group. In contrast, the experimental groups all showed learning gains between the spoken/written form recognition pretest, posttest, and delayed posttest.

The analysis showed that the assumption of sphericity had been violated for C ($\chi^2(2) = 0.00$, $p = 0.00$), LL($\chi^2(2) = 10.259$, $p = 0.006$), and LSL ($\chi^2(2) = 37.536$, $p < 0.001$) groups; therefore, a Greenhouse-Geisser correction was used. The within-participant main effect (time) indicated no significant results for C ($F(1.00, 20.00) = 2.105$, $p = 0.162$). In contrast, the results for LL ($F(1.294, 18.114) = 192.063$, $p < 0.001$ with partial $\eta^2 = 0.932$), and LSL ($F(1.066, 20.259) = 242.106$, $p < 0.001$ with partial $\eta^2 = 0.927$), were both statistically significant with

very large effect sizes. As for L ($\chi^2(2) = 1.289$, $p = 0.525$) and LS ($\chi^2(2) = 4.799$, $p = 0.091$), the assumption of sphericity was met. Furthermore, the within-participant main effect (time) was statistically significant for both groups with very large effect sizes; for L, $F(2, 40) = 420.0$, $p < 0.001$ with partial $\eta^2 = 0.955$, and for LS, $F(2, 40) = 514.126$, $p < 0.001$ with partial $\eta^2 = 0.963$.

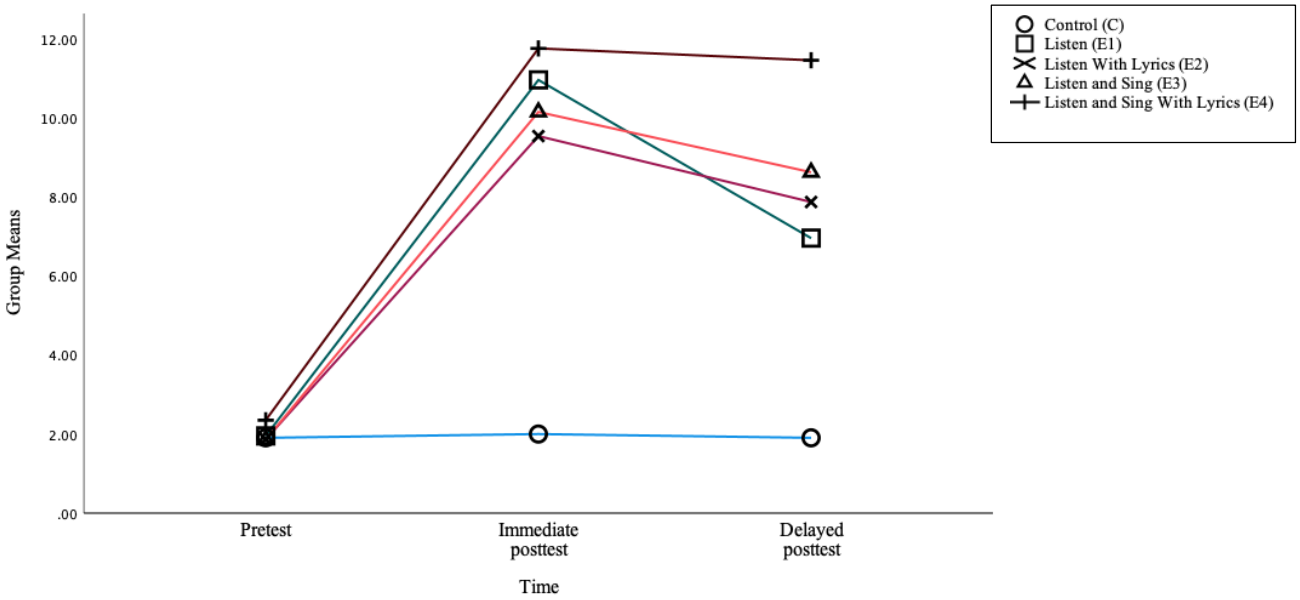


Figure 3.2 Group means for receptive vocabulary learning over time.

3.7.1.2 Measuring Productive Knowledge of FS

Figure 3.3 shows that no learning gains were made between the written form recall pretest, posttest, and delayed posttest for the control (C) group. In contrast, the experimental groups all showed learning gains between the form recall pretest, posttest, and delayed posttest.

The analysis showed that the assumption of sphericity had been violated for all groups; C ($\chi^2(2) = 8.78$, $p = 0.012$), L ($\chi^2(2) = 10.540$, $p < 0.005$), LL ($\chi^2(2) = 9.761$, $p = 0.008$), LS ($\chi^2(2) = 27.484$, $p < 0.001$), LSL ($\chi^2(2) = 39.105$, $p < 0.001$), and therefore, a Greenhouse-Geisser correction was used. The results indicated that the C group's within-participant main effect (time)

was not statistically significant $F(1.460, 29.197) = 1.00$, $p = 0.357$. As for the experimental groups, they all showed statistically significant within-participant main effect (time) with very large effect sizes; L, $F(1.403, 28.015) = 112.085$, $p < 0.001$ with partial $\eta^2 = 0.849$, LL, $F(1.309, 18.324) = 128.469$, $p < 0.001$ with partial $\eta^2 = 0.902$, LS, $F(1.133, 22.668) = 99.457$, $p < 0.001$ with partial $\eta^2 = 0.833$, and LSL, $F(1.060, 20.147) = 165.005$, $p < 0.001$ with partial $\eta^2 = 0.897$.

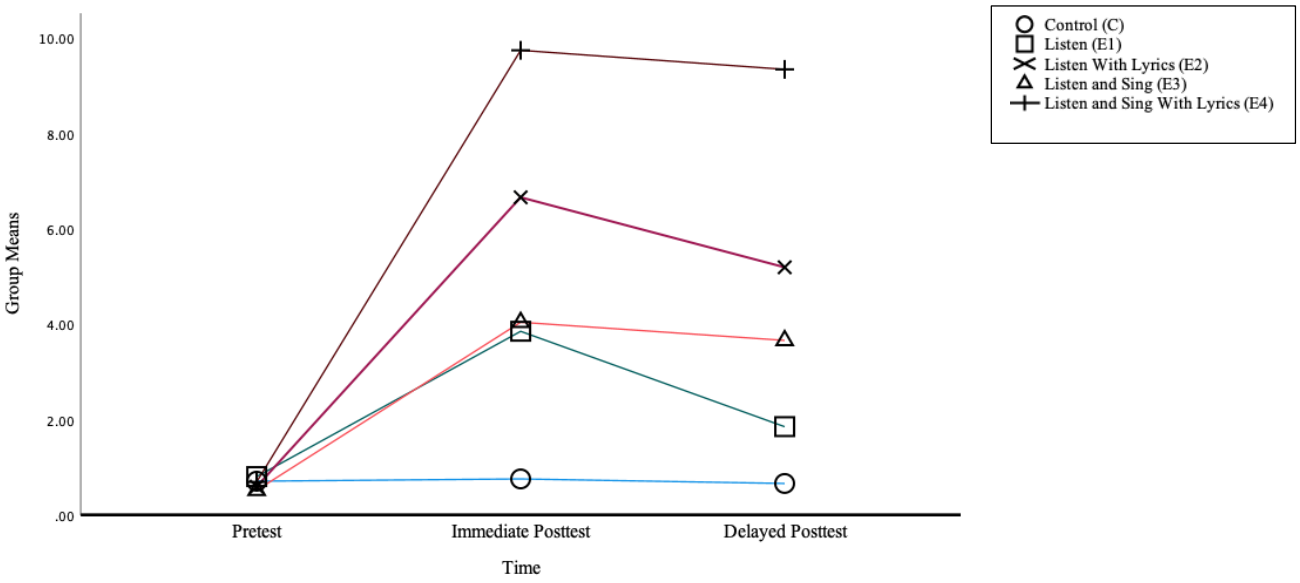


Figure 3.3 Group means for productive vocabulary learning over time.

3.7.2 FS Learning Gains

The pairwise comparison between test times (pretest, immediate posttest, and delayed posttest) was used to determine the receptive, productive, and total FS learning gains under each listening condition.

The results showed no significant receptive or productive FS learning gains for the control (C) group. The different listening conditions for the experimental groups showed significant gains with very large ($0.9 < d < 2.0$) to extremely large ($d > 2.0$) effect sizes. See further details on the pairwise comparison scores of the pretest, immediate posttest, and delayed posttest in Table 2.6.

Table 2.6 Pairwise comparison for different sections of the tests

Time of testing (*i*)	Time of testing(*j*)	Difference between means (*j*-*i*)	*SD* error	*p*	Cohen's *d*	95% confidence interval for the difference	
						Upper bound	Lower bound
C Productive							
1	2	0.048	0.048	0.988	0.057	0.172	-0.077
	3	-0.048	0.084	1.000	0.064	0.171	-0.267
C Receptive							
1	2	0.095	0.066	0.487	0.088	0.267	-0.076
	3	0.000	0.000	-	0.000	0.00	0.00
C Total							
1	2	0.190	0.088	0.127	0.127	0.420	-0.039
	3	0.000	0.098	1.000	0.000	0.255	-0.255
L Productive							
1	2	3.048*	0.253	<0.001	2.788	3.709	2.386
	3	1.048*	0.129	<0.001	1.077	1.384	0.711
L Receptive							
1	2	9.000*	0.309	<0.001	6.071	9.806	8.194
	3	5.000*	0.345	<0.001	3.495	5.901	4.099
L Total							
1	2	12.048*	0.514	<0.001	5.078	13.390	10.705
	3	6.048*	0.399	<0.001	2.875	7.091	5.005
LL Productive							
1	2	6.067*	0.502	<0.001	3.480	8.685	6.648
	3	4.600*	0.400	<0.001	3.144	7.047	4.953
LL Receptive		7.667*					
1	2	6.000*	0.475	<0.001	4.385	8.685	6.648
	3	13.733*	0.488	<0.001	3.247	7.047	4.953
LL Total							
1	2	13.733*	0.859	<0.001	4.153	15.576	11.891
	3	10.600*	0.767	<0.001	3.396	12.245	8.955
LS Productive							
1	2	3.524*	0.349	<0.001	3.396	4.436	2.611
	3	3.143*	0.303	<0.001	2.750	3.935	2.351
LS Receptive							
1	2	8.286*	0.317	<0.001	5.160	9.115	7.457
	3	6.762*	0.292	<0.001	4.620	7.525	5.999
LS Total							
1	2	11.810*	0.627	<0.001	4.340	13.448	10.171
	3	9.905*	0.502	<0.001	4.155	11.216	8.593
LSL Productive							
1	2	9.050*	0.709	<0.001	4.147	10.911	7.189
	3	8.650*	0.653	<0.001	4.250	10.364	6.936

LSL Receptive							
1	2	9.400*	0.604	<0.001	4.927	10.987	7.813
	3	9.100*	0.571	<0.001	5.133	10.598	7.602
LSL Total							
1	2	18.450*	1.221	<0.001	4.910	21.656	15.244
	3	17.750*	1.140	<0.001	5.050	20.744	14.756

Note. C= Control, L= Listen only, LL= Listen while reading the lyrics, LS= Listen and sing, LSL= Listen and sing while reading the lyrics

3.7.3 Receptive VS Productive Song Use

In response to the second research question, comparing the effects of receptive and productive use of songs on incidental FS learning, Multivariate analysis of variance (MANOVA) was used.

3.7.3.1 Immediate Posttest

The analysis revealed an overall statistically significant difference, F (8.00, 186.00) = 70.32, p < 0.001, Pillais' Trace = 1.502, partial $\eta^2 = 0.751$ on the immediate posttest. The results for different sections of the test considered separately showed significant differences across sections. Both receptive recognition (F (4, 93) = 86.819; $p < 0.001$; partial $\eta^2 = 0.732$) and productive recall (F (4, 93) = 63.510; $p < 0.001$; partial $\eta^2 = 0.751$) showed extremely large effect sizes.

To determine whether there's a relationship between the type of treatment and performance on each section of the test, the Post hoc comparison using the Bonferroni test was used. Based on the previous section of this study, it is clear that no learning occurred as a result of the control group; therefore this section will focus on the differences between the two experimental groups, For the receptive recognition of FS, the only significant difference was between the LL and LSL groups, with LSL outperforming LL ($p = 0.011$, $d = 0.981$) with a very large effect size.

On the productive recall section of the test, statistically significant differences were found between all but one comparison (i.e., L & LS, not significant). The comparison between the rest

of the experimental groups is as follows: L and LL ($p < 0.001$, $d = 1.496$) showed LL outperforming L with a very large effect size. L and LSL ($p < 0.001$, $d = 2.641$) showed LSL outperforming L with an extremely large effect size. LL and LS ($p < 0.001$, $d = 1.291$) showed LL outperforming LS with a very large effect size. LL and LSL ($p < 0.001$, $d = 1.169$) showed LSL outperforming LL with a very large effect size. Finally, the comparison between LS and LSL ($p < 0.001$, $d = 2.423$) showed LSL outperforming LS with an extremely large effect size.

3.7.3.2 Delayed Posttest

The analysis revealed an overall statistically significant difference, F (8.00, 186.00) = 48.543, $p < 0.001$, Pillais' Trace = 1.352, partial $\eta^2 = 0.676$ on the delayed posttest. The results for different sections of the test considered separately showed significant differences across sections. Both receptive recognition (F (4, 93) = 73.922; $p < 0.001$; partial $\eta^2 = 0.761$) and productive recall (F (4, 93) = 81.293; $p < 0.001$; partial $\eta^2 = 0.778$) showed extremely large effect sizes.

To determine whether there's a relationship between the type of treatment and performance on each section of the test, the Post hoc comparison using the Bonferroni test was used. Based on the previous section of this study, it is clear that no learning occurred as a result of the control group; therefore, this section will focus on the differences between the experimental groups, For the receptive recognition of FS, the comparison of the receptive recognition section of the test revealed that LS ($p = 0.042$, $d = 0.935$), and LSL ($p < 0.001$, $d = 2.230$) both outperformed L with very large and extremely large effect sizes. LSL also outperformed LL ($p < 0.001$, $d = 1.61$) and LS ($p < 0.001$, $d = 1.39$) with very large effect sizes.

For the productive recall section of the test, the analysis revealed that LL ($p < 0.001$, $d = 2.164$), LS ($p = 0.008$, $d = 1.457$), and LSL ($p < 0.001$, $d = 3.698$) all outperformed the L group

with very large to extremely large effect sizes. Furthermore, LSL outperformed the LL ($p < 0.001$, $d = 1.77$) and LS ($p < 0.001$, $d = 2.6511$) with very large to extremely large effect sizes.

No significant difference was found between the LL and LS groups on either section of the delayed posttest. In addition, no statistically significant difference was found between L and LL on the receptive section of the delayed posttest.

3.8 Discussion

3.8.1 Does spaced listening to a song contribute to L2 formulaic sequences learning?

In response to the first research question investigating incidental learning of formulaic sequences from receptive and productive use of songs, the results showed that repeated exposure (i.e., Five times) involving spaced practice has the potential to foster FS learning. The learning gains of the participants in the experimental groups are as follows; In the listening-only group, learning gains of 37.66 % (12.05 items) from the pretest to the immediate posttest and retained knowledge of 18.91% (6.05 items) on the delayed posttest. In the listening and reading the lyrics group, learning gains of 42.91% (13.73 items) from the pretest to the immediate posttest and retained knowledge of 33.13% (10.60 items) on the delayed posttest. In the listening and singing group, learning gains of 36.91% (11.81 items) from the pretest to the immediate posttest and retained knowledge of 30.97% (9.91 items) on the delayed posttest. Lastly, in listening and singing while reading the lyrics group, participants showed learning gains of 57.66% (18.45 items) from the pretest to the immediate posttest and retained knowledge of 55.47% (17.75 items) on the delayed posttest. In contrast, participants in the control group did not show any statistically significant learning gains from the pretest to the immediate posttest and delayed posttest, with gains of 0.59% (0.19 items) from the pretest to the immediate posttest and no learning gains on the delayed posttest.

Measuring the potential learning of receptive and productive knowledge of FS separately, the results of the form-recognition test revealed no learning for the control group from the pretest to the immediate posttest and the delayed posttest. All the different treatments in the experimental groups showed statistically significant learning gains on the immediate posttest, which were as follows; L, 56.25% (9 items), LL, 47.88% (7.66 items), LS, 51.75% (8.28 items), and LSL, 58.75% (9.4 items). On the form-recognition delayed posttest, the experimental groups retained a significant portion of the learning gains: L, 31.25% (5 items), LL, 37.50% (6 items), LS, 42.25% (6.76 items), and LSL, 56.88 (9.1 items).

The results on the written form recall test results revealed no learning for the control group from the pretest to the immediate posttest and the delayed posttest. All the different treatments in the experimental groups showed statistically significant learning gains on the immediate posttest, which were as follows; L, 19.06% (3.05 items), LL, 37.75% (6.04 items), LS, 22.06% (3.53 items), and LSL, 56.56% (9.05 items). On the written form recall delayed posttest, the experimental groups retained a significant portion of the learning gains: L, 6.56% (1.05 items), LL, 28.75% (4.6 items), LS, 19.69% (3.15 items), and LSL, 54.06% (8.65 items).

The current study shows greater learning gains in incidental learning of FS from various aural inputs compared to previous research. Pavia et al. (2019) found no significant gains from listening to one song and learning gains of 10.97% for collocation recognition from listening to another song. The difference in proficiency levels between the participants in the two studies may account for the variance. Majuddin et al. (2021) found that higher scores on the VST, which was taken as a measure of participants' proficiency, led to a higher likelihood of learning multiword expressions from audiovisual materials. Therefore, it is expected to find larger learning gains in the current study with more proficient participants (i.e., knowledge of around

1,593 of the most frequent word families) compared to the participants in Pavia et al. (2019) (i.e., knowledge of around 430 of the most frequent word families).

The results are slightly higher than those in the reading-while-listening study by Webb and Chang (2022). Their best learning condition, reading while listening to a graded reader, resulted in learning gains of 28.46% (4.64 collocations). This can be due to the number of exposures to the target collocations (i.e., 1-16 encounters). Majuddin et al. (2021) found that additional exposure to multiword expressions through audiovisual material could increase the likelihood of learning by 2.69 times after one additional viewing. With that in mind, in the current study, the number of encounters after five listening sessions ranged from 5-30, much greater than the number of encounters in the Webb & Chang (2022) study, with 1-16 encounters. Further to this point, Webb et al. (2013), investigating incidental learning of collocations from reading while listening to a graded reader, found learning gains of 27% (3.06 collocations), 33% (3.66 collocations), 55% (6.53 collocations), and 76% (8.24 collocations) from 1, 5, 10, and 15 encounters.

Webb et al.'s (2013) results on learning the form of collocations seem more consistent with the results found in the current study. However, it is important to note that in the 15-encounter group of the Webb et al. study, every collocation was encountered 15 times, while in the current study, 50% of the FS were encountered five times, 37.5% were encountered ten times, and only 6.25% were encountered between 20-30 times. This suggests that incidental learning of FS from songs can occur with fewer encounters in comparison to incidental learning from reading while listening to graded texts. This may be due to the importance of speech prosody in noticing and acquiring FS. In a review of empirical studies in L1 and L2 language development, Lin (2019) argued that prosodic patterns could prompt learning formulaic

sequences (FS) by dividing the input into more easily discernable lexical units. The more prosodically salient the source of input is, the higher the chance of FS being learned. For example, the vocal phrasing of FS, such as “let it go” and “running in circles” emphasized these chunks in the current study. Similar patterns were found in Tomczak and Lew’s (2019) study. They showed that participants retained multi-word units better in the song-based deliberate teaching condition than in the spoken-sentence deliberate teaching condition. Although it was not explicitly stated, upon closer examination of their instruments and methodology, it appears that the multi-word targets were more prosodic in the song treatment conditions. For example, “part-time lover” is highly emphasized through vocal phrasing and, therefore, more prosodic. This suggests that multiword units can be highly salient in lyric-based sources of input, which can make songs a helpful source of language input in particular for FS learning and teaching (Werner, 2020).

Lastly, comparing the learning gains on the receptive test compared to the productive test, participants showed greater receptive learning gains than productive learning gains. This suggests that productive knowledge is more difficult to acquire than receptive knowledge of words from songs. This pattern of acquisition is consistent with results found in other incidental reading and listening studies investigating incidental learning of multiword units (Webb et al., 2013) and single-word units (Teng, 2016; Webb, 2005, 2007, 2009).

3.8.2 Which receptive/productive use of a song leads to better FS learning?

In response to the final research question comparing the learning gains between different treatments, the results showed statistically significant differences between LL and LSL groups on the form-recognition test, with LSL outperforming LL by approximately 14% (2.217 items). This was the only significant difference between the experimental groups on the receptive aspect of

word knowledge. This lack of variance between groups can be due to the ceiling effect on the form recognition test. For example, the majority (*M*=96%) of the participants across experimental groups selected the correct pairing for the target item “run,” which was “away.”

The results showed a statistically significant difference between all but one paring, LL and LS, on the more difficult test, form recall. LSL outperformed all groups, followed by the LL group, which outperformed the L and LS groups. This is not surprising, as the LSL and LL both had the written form of the target words present during the treatment period, so they were better able to recall and produce the written form of the target words on the immediate and delayed posttests. Previous research (Lotto & de Groot, 1998; Majuddin et al., 2021; Schmitt, 2010) suggests that congruency between learning conditions and testing methods leads to better performance. However, considering the fact that LSL also outperformed LL, the learning gains cannot solely be attributed to the congruency effect between treatment and testing conditions. The combination of singing with access to the lyrics in the LSL condition increased the quality of engagement with the source of input, the song, resulting in better learning outcomes. According to González-Fernández (2022), the primary cause of poor learning outcomes in listening activities that involve songs could be due to the low quality of engagement. A previous study by González-Fernández and Schmitt (2015) revealed that the quality of engagement is a better predictor of collocation knowledge (*r*=0.45) than years of study, suggesting that large quantities of input alone are not enough for durable learning of multiword units. This can explain the lack of correlation between large quantities of input and vocabulary knowledge from listening to songs reported in Peters's (2020) study. In their study, like the listening-only condition of the current study, they only reported passive exposure to songs, which is not the optimal listening condition for language learning to occur. In conclusion, listening and singing

while reading the lyrics was the most effective learning condition for learning both receptive and productive aspects of FS from the target song, which is consistent with previous research.

3.9 Conclusion and Future Direction for Research

The results of the present study indicated that formulaic sequences could be learned from songs incidentally. Additional factors, including frequency of encounters and prior vocabulary knowledge, can influence learning gains. Engaging in singing activities and using song lyrics can boost engagement quality and increase the chances of incidental learning. This makes songs a valuable input source for those learning a second or foreign language. Nevertheless, several limitations to this study may provide a path for future research. First, the number of encounters between different target FS was different. For example, the FS "run away" was encountered 30 times compared to "turn around," which only occurred five times after the five listening sessions. This may have contributed to the ceiling effect on the earlier receptive knowledge test, making learning more frequent FS items easier for participants across treatment conditions. While using authentic songs increases the ecological validity of the study, it may be useful to explore the effect of encounters with a specific FS item in addition to different listening conditions. To achieve this, one can create a song that eliminates any potential effects from differences between FS items and controls for the frequency of encounter effect on learning.

Another limitation is at the target FS level. This study did not control for the L1-L2 congruency effect (i.e., having a similar L1 translation) between target FS items. Previous studies investigating the effects of L1-L2 collocation congruency have produced mixed results. Peters (2016), investigating the effects of congruency and word length on deliberate learning of L2 collocations, found incongruent collocations more difficult for L2 learners than congruent collocations. The same patterns were found in Nguyen and Webb (2017), Wolter and Gyllstad

(2011, 2013), Yamashita and Jiang (2010), and Wolter and Yamashita (2018), suggesting that congruency could positively affect the learning of L2 FS and that second/foreign language learners process congruent FS faster and more accurately than non-congruent FS. In contrast, Puimège and Peters (2020) found no significant effect for congruency on the incidental learning of L2 collocations from viewing television. Thus, it would be useful to investigate the effects of L1-L2 congruency on incidental learning of FS from songs.

Finally, the current study solely focused on the learning of FS form through receptive recognition and productive recall. Measuring knowledge of meaning was not considered, so it is not clear whether the participants understood what the target FS meant. The knowledge of form is the first step to learning FS (González-Fernández & Schmitt, 2020), and it may prime learners to recognize the target FS in future encounters in other texts with more informative contexts. Nevertheless, further research is needed to understand better how the connection between form and meaning of FS is made through learning from songs.

This study is one of only a few to have empirically examined the impact of using songs for language learning. Therefore, there are still many approaches to using songs for language learning that teachers have reported anecdotally (Arnold & Herrick, 2017) but which have yet to be studied in an empirical context. Research has demonstrated that retrieval can improve vocabulary learning, as evidenced by studies on word learning (Boers, 2021). On the other hand, teachers have observed that students can easily recall song lyrics from memory (Werner, 2018). Therefore, it would be beneficial to conduct empirical research on the effects of retrieval through listening to songs on vocabulary learning to connect research with practical applications.

3.11 References

Arnold, J. L., & Herrick, E. (2017). *New ways in teaching with music.* TESOL International Association. Alexandria, Virginia, USA: TESOL press.

Beran, E. (2000). Review of Oakes, M. P. (1998). Statistics for Corpus Linguistics. *International Journal of Applied Linguistics, 10*(2), 269-274. doi: 10.1111/j.1473-4192.2000.tb00152.x

Boers, F. (2021). *Evaluating second language vocabulary and grammar instruction: A synthesis of the research on teaching words, phrases, and pattern*s. Routledge

Boers, F, Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research, 10*: 245–261. https://doi.org/10.1191/1362168806lr195oa

Boers, F., & Webb, S. (2018). Teaching and learning collocation in adult second and foreign language learning. *Language Teaching, 51*(1), 77–89. https://doi.org/10.1017/S0261444817000301

Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental Vocabulary Acquisition from Reading, Reading-While-Listening, and Listening to Stories. Reading in a Foreign Language, 20(2), 136–163.

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text & Talk, 20*(1), 29–62. https://doi.org/10.1515/text.1.2000.20.1.29

Gary, J., & Gary, N. (1981). Caution: Talking may be dangerous to your linguistic health. The case for much greater emphasis on listening comprehension in foreign language instruction. *International Review of Applied Linguistics in Language Teaching, IRAL, 19*(1), 1–14. https://doi.org/10.1515/iral.1981.19.1-4.1

González Fernández, B. (2022). Vocabulary knowledge: Lexical depth and its relationship with out-of-class exposure. In P. Szudarski, & S. Barclay (Eds.), *Vocabulary Theory, Patterning and Teaching* (pp. 21-47). SLA

González Fernández, B., & Schmitt, N. (2015). How much collocation knowledge do L2 learners have?: the effects of frequency and amount of exposure. ITL, 166(1), https://doi.org/10.1075/itl.166.1.03fer

González-Fernández, B., & Schmitt, N. (2020). Word knowledge: Exploring the relationships and order of acquisition of vocabulary knowledge components. *Applied Linguistics, 41*(4), 481–505. https://doi.org/10.1093/applin/amy057

Guglielmino, L. (1986). The affective edge: Using songs and music in ESL instruction. Adult Literacy and Basic Education,10, 19–26.

Hill, J. (2001). Revising priorities: From grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching collocation: Further development in the lexical approach* (pp. 47–69). Oxford University Press.

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125-132.

Howarth, P. (1998). Phraseology and Second Language Proficiency. *Applied Linguistics, 19*(1), 24–44. https://doi.org/10.1093/applin/19.1.24

Kim, S., & Webb, S. (2022). The Effects of Spaced Practice on Second Language Learning: A Meta-Analysis. *Language Learning*, *72*(1), 269–319. https://doi.org/10.1111/lang.12479

Laufer, B. (1998). The development of passive and active vocabulary: Same or different? *Applied Linguistics, 19*(2), 255–271. https://doi.org/10.1093/applin/19.2.255

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, *61*(2), 647–672. https://doi.org/10.1111/j.1467-9922.2010.00621.x

Lin, P. M. (2012). Sound evidence: The missing piece of the jigsaw in formulaic language research. *Applied Linguistics, 33*, 342–347.

Lin, P. (2019). *Formulaic language and speech prosody.* In A. Siyanova-Chanturia, & A. Pellicer-Sánchez (Eds.), *Understanding Formulaic Language A second Language Acquisition Perspective* (pp 78-94). Routledge

Lindgren, E., & Muñoz, C. (2013). The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism,10*, 105–129. https://doi.org/10.1080/14790718.2012.679275

Lock, S. (2019). *Units sales of karaoke machines in the U.S.* 2008-2018. Statista. Retrieved from https://www.statista.com/statistics/454634/number-of-karaoke-machines-in-the-us/.

Lotto, L., & de Groot, A. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning, 48*(1), 31–69. https://doi.org/10.1111/1467-9922.00032

Laufer, B. (2003). Vocabulary Acquisition in a Second Language: Do Learners Really Acquire Most Vocabulary by Reading? Some Empirical Evidence. Canadian Modern Language Review, 59(4), 567–587. https://doi.org/10.3138/cmlr.59.4.567

Macis, Sonbul, S., & Alharbi, R. (2021). The effect of spacing on incidental and deliberate learning of L2 collocations. System (Linköping), 103, 102649- https://doi.org/10.1016/j.system.2021.102649

Majuddin, E., Siyanova-Chanturia, A., & Boers, F. (2021). Incidental Acquisition of Multiword Expressions through Audiovisual Materials: The Role of Repetition and Typographic Enhancement. *Studies in Second Language Acquisition, 43*(5), 985–1008. https://doi.org/10.1017/S0272263121000036

McCauley, S. M., & Christiansen, M. H. (2017). Computational investigations of multiword chunks in language learning. *Topics in Cognitive Science, 9*(3), 637–652.

Medina, S. L. (1993). The effects of music on second language vocabulary acquisition. *National Network for Early Language Learning, 6*(3), 1-8.

Morgan, B. Q., & Oberdeck, L. M. (1930). Active and passive vocabulary. In E. W. Bagster-Collins (Ed.), *Studies in modern language teaching* (pp. 213–221). Basingstoke: Macmillan.

Muijs, D. (2011). *Doing quantitative research in education with SPSS*. London, UK: SAGE Publishers.

Nation, I. S. P. (2013). *Learning vocabulary in another language (2nd ed.).* New York, NY: Cambridge University Press.

Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Boston: Heinle.

Nelson, R. (2018). How “chunky” is language? Some estimates based on Sinclair’s Idiom Principle. *Corpora, 13*(3), 431–460. https://doi.org/10.3366/cor.2018.0156

Newton, J. M., & Nation, I. S. P. (2020). *Teaching ESL/EFL listening and speaking*. Routledge

Nguyen, T. M. H., & Webb, S. (2017). Examining second language receptive knowledge of collocation and factors that affect learning. Language Teaching Research : LTR, 21(3), 298–320. https://doi.org/10.1177/1362168816639619

Pavia, N., Webb, S., & Faez, F. (2019). Incidental vocabulary learning through listening to songs. *Studies in Second Language Acquisition*, 1-24. doi:10.1017/S0272263119000020

Peters, E. (2016). The learning burden of collocations: The role of interlexical and intralexical factors. Language Teaching Research : LTR, 20(1), 113–138. https://doi.org/10.1177/1362168814568131

Peters, E. (2020) The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. In S. Webb (Ed), Approaches to Learning, Testing and Researching L2 Vocabulary (Vol. 109). John Benjamins Publishing Company.

Peters, E., Heynen, E., & Puimège, E. (2016). Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System (Linköping), 63*, 134–148. https://doi.org/10.1016/j.system.2016.10.002

Puimège, E., & Peters, E. (2020). Learning formulaic sequences through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, *42*(3), 525–549. https://doi.org/10.1017/S027226311900055X

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan.

Sinclair, J. (1991). *Corpus, concordance, collocation.* Oxford: Oxford University Press.

Siyanova-Chanturia, A., & Pellicer-Sánchez, A. (2019). *Understanding formulaic language: a second language acquisition perspective*. Routledge.

Swain, M. (1985). Communicative competence: some roles of comprehension input and output in its development. In S. Gass and C. Madden (Eds), *Input in Second Language Acquisition* (pp. 235-253). Rowley, Mass: Newbury House.

Swains, M. (2005). The output hypothesis: theory and research. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (pp. 471-483). Mahwah, N.J.: Lawrence Erlbaum Association.

Tegge, F. (2015). *Investigating song-based language teaching and its effect on lexical learning* (Unpublished doctoral dissertation). Victoria University of Wellington, New Zealand. doi: 10.13140/RG.2.1.3159.8560.

Tegge, F. (2017). The lexical coverage of popular songs in English language teaching. *System, 67*, 87–98. doi https://doi.org/10.1016/j.system.2017.04.016.

Tegge, F. (2018). Pop songs in the classroom: Time-filler or teaching tool? *ELT Journal, 72*, 274–284. doi: https://doi.org/10.1093/elt/ccx071.

Teng, F. (2016). Incidental vocabulary acquisition from reading-only and reading-while listening: A multi-dimensional approach. *Innovation in Language Learning and Teaching,12*, 274–288. https://doi.org/10.1080/17501229.2016.1203328

Tomczak, E., & Lew, R. (2019). "The Song of Words": Teaching multi-word units with songs. *3L: The Southeast Asian Journal of English Language Studies, 25*(4), 16–33.

Van Zeeland, H.,&Schmitt, N. (2013). Lexical coverage and L1 and L2 listening comprehension: The same or different from reading comprehension. *Applied Linguistics, 34,* 457–479. doi: 10.1093/applin/ams074

Warner. R. M. (2013). Applied statistics: from bivariate through multivariate techniques (2nd ed.). *SAGE Publications.*

Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition*, *27*(1), 33-52. https://doi.org/10.1017/S0272263105050023

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, *28*(1), 46-65. https://doi.org/10.1093/applin/aml048

Webb, S. (2008). Receptive and productive vocabulary size. *Studies in Second Language Acquisition*, *30*(1), 79-95. https://doi.org/10.1017/S0272263108080042

Webb, S. (2009). The effects of receptive and productive learning of word pairs on vocabulary knowledge. *RELC Journal, 40*(3), 360-376.

Webb, S. A., & Chang, A. C-S. (2012). Vocabulary learning through assisted and unassisted repeated reading. *Canadian Modern Language Review, 68*, 267–290. doi 10.3138/cmlr.1204.1

Webb, S., & Chang, A. C.-S. (2022). How does mode of input affect the incidental learning of collocations? Studies in Second Language Acquisition, 44(1), 35–56. https://doi.org/10.1017/S0272263120000297

Webb, S., & Nation, P. (2017). *How vocabulary is learned.* Oxford, UK: Oxford University Press.

Webb, S., Newton, J., & Chang, A. C-S. (2013). Incidental learning of collocation. *Language Learning*, 63(1), 91-120. https://doi.org/10.1111/j.1467-9922.2012.00729.x

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. ITL. *International Journal of Applied Linguistics, 168*(1), 34-70.

Wray, A. (2000). Formulaic sequences in second language teaching: principle and practice. *Applied Linguistics, 21*(4), 463–489. https://doi.org/10.1093/applin/21.4.463

Werner, R. (2018). Music, movement and memory: Pedagogical songs and mnemonic aids. TESOL Journal, 9(4), 1–. https://doi.org/10.1002/tesj.387

Werner, R. (2020). “Song-advantage” or “Cost of Singing”? : A Research Synthesis of Classroom-based Intervention Studies Applying Lyrics-based Language Teaching (1972–2019). *Journal of Second Language Teaching & Research, 8* (1), 138-170. DOI: 10.20378/irb-49313

Wolter, B., & Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics, 32*, 430–449.

Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing. Studies in Second Language Acquisition, 35, 451–482.

Wolter, B., & Yamashita, J. (2018). Word Frequency, Collocational Frequency, L1 Congruency, and Proficiency in L2 Collocational Processing: What Accounts for L2 Performance? *Studies in Second Language Acquisition*, *40*(2), 395–416. https://doi.org/10.1017/S0272263117000237

Yamashita, J., & Jiang, N. (2010). L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. TESOL Quarterly, 44, 647–668.

Zhou, X., & Tarocco, F. (2007). *Karaoke: The global phenomenon.* London, UK: Reaktion Books Ltd.

Chapter 4: Study 3

The Effects of Retrieval on Incidental Vocabulary Learning from Songs

4.1 Abstract

This study examines the effectiveness of retrieval implemented in the input-input-output-output-input sequence of activities to acquire new formulaic sequences (FS) from a song. Songs have been shown to foster language acquisition (Good et al., 2015; Pavia et al., 2019; Tegge, 2015), especially with assisted use of lyrics and sing-along activities (See Article Two in this thesis). However, the effects of retrieval and sequential learning which have been shown to affect vocabulary learning through listening and speaking tasks, deliberate learning of word pairs, and other visual sources of input, such as TED talks (de la Fuente, 2002; Nakata, 2017; Nakata et al., 2020; Nguyen & Boers, 2019) have yet to be explored with song use for foreign language acquisition. In this study, Thai students (N=60) learning English as a foreign language (EFL) were randomly assigned to control, non-retrieval, and retrieval groups. The song previously used in chapter three of this thesis was used. The control group (C) participants completed the pretest, immediate posttest, and two weeks delayed posttest without any exposure to the target language presented in song format. The non-retrieval group (E1) Listened, listened with lyrics, sang along with lyrics, sang along without lyrics, and listened only one last time. The retrieval group (E2) listened, listened with lyrics, sang along with lyrics, wrote down the lyrics from memory while listening to the instrumental version of the song, and finally, listened only one more time. The results indicated that the two experimental groups outperformed the control group significantly on the immediate and delayed posttests. In addition, the retrieval group (E2) outperformed the non-retrieval group (E1) significantly on both the immediate and delayed posttests, with a very large effect size.

Keywords: Vocabulary learning; song; retrieval; sequential practice; multimodal input

4.2 Introduction

Learning a second/foreign language can be challenging. Research has shown that vocabulary knowledge is one of the main predictors of proficiency (Qian, 2002; Schmitt, 2008). It is estimated that learners need to master anywhere between 3,000 to 9,000 most frequent word families in English to engage in different activities, including conversational English, reading books, and watching television (Nation, 2013). The depth of vocabulary knowledge adds another layer of difficulty to acquiring this large quantity of words. For instance, having a deep understanding of a word would mean knowing its form, meaning, and use (Nation, 2013). The

limited time in language classrooms and the lack of motivation or direction in learners outside the classroom have led researchers to decades-long creation and evaluation of different vocabulary learning activities. As vocabulary activities have developed, so have theories and frameworks for evaluating their effectiveness (Hulstijn & Laufer, 2001). One of the models of language acquisition and by proxy vocabulary learning is proposed by Nation (2007), the *Four Strands.* This model used the time-on-task principle to propose that learners need to spend an equal amount of time on meaning-focused input, meaning focused output, language-focused learning, and fluency development activities. Considering that three of the four strands comprise meaning-focused activities, this model highlights the importance of engagement in meaning-focused activities in the development of communicative competence. The language skills acquired through meaning-focused activities are largely due to incidental learning (Nation, 2007). Incidental learning refers to the process of acquiring language skills as a by-product of an activity where learners are primarily focused on the message content rather than precisely how this content is expressed (Boers, 2021). Some examples of incidental learning activities are reading novels, listening to stories, watching videos such as TED talks and television series, and listening to songs (Horst, 2005; Jin & Webb, 2020; Nguyen & Boers, 2019; Puimège & Peters, 2020; Pavia et al., 2019). Songs and lyric-based learning have gained popularity among educators and researchers in recent years (Werner, 2020; Pavia et al., 2019). Recent reports suggest that songs are one of the most used forms of language exposure outside the classroom (Lindgren & Munoz, 2013; Lai et al., 2015; Peters, 2020) and are frequently used by teachers inside the classrooms (Tegge, 2015; Werner, 2020). Nevertheless, songs remain one of the least empirically researched sources of input for incidental learning. This article explores vocabulary learning through song exposure and use in a classroom setting. The participants were neither told

about vocabulary learning nor forewarned of a vocabulary test to ensure the conditions of the current study were consistent with the characteristics defining incidental learning (Hulstijn, 2001).

Previous studies have speculated that the prosodic nature of songs makes formulaic sequences (FS) highly salient and contribute to FS learning (Pavia et al., 2019; Tomczak &Lew, 2019; Pavia, chapter 2&3 of this volume). Considering the importance of FS for receptive and productive language use and language learners' difficulty acquiring FS (Boers, 2020), incidental learning of FS was used to measure vocabulary learning.

Furthermore, studies on language acquisition through audiovisual sources of input have demonstrated that the process of input-output-input enables learners to identify gaps in their vocabulary knowledge during the output session, thereby increasing their likelihood of learning those vocabulary items during the subsequent input session (Nguyen & Boers, 2019). For this study, the above-mentioned cycle was adapted to include five sessions (i.e., input-input-output-output-input) since previous research on using songs for incidental vocabulary learning has shown that five repeated exposures to the target song led to higher learning levels than just three sessions (Pavia et al., 2019).

Finally, the effects of attention-raising practices through retrieval on vocabulary learning are explored. In a survey, Peters (2020) found a poor relationship between exposure to songs and overall vocabulary knowledge despite large quantities of exposure to songs in out-of-school leisure activities. González Fernández (2022) suggests that this poor correlation between the quantity of exposure and vocabulary learning in out-of-school song-based activities, predominantly listening, may be due to the lack of quality of engagement. For learning to take place, the quantity of exposure alone does not suffice; there must be a degree of attention and

involvement with the activity and its underlying language for learning to occur (Webb & Nation, 2017; González Fernández & Schmitt, 2015; Laufer & Hulstijn, 2001; Yanagisawa & Webb, 2021; 2022). Consequently, to increase the quality of engagement during song-based activities, the use of retrieval is explored. Webb and Nation (2017) identify retrieval as a powerful learning tool that contributes specifically to the "quality of attention." (p.69).

4.3 Literature Review

4.3.1 Vocabulary Acquisition from Songs

In 2017 Arnold and Herrick (Eds) compiled a comprehensive text offering 110 ways of using music and songs for language education. This book, "*The New Ways in Teaching with Music,"* responded to the gap that while it is accepted that music is used for language education by teachers and students, no clear instructions or reports were available as a guide for implementation and use (Engh, 2013). Therefore, this book provided examples of activities reported by students, educators, teachers etc., for using songs in the classroom, making it an invaluable resource for anyone involved in language education. Nevertheless, the lack of empirical evidence supporting the practices reported in the previously mentioned book and other how-to guides for using songs (Murphey, 1992a) has made it difficult to understand how effective songs are for language education. Empirical studies allow deliberate planning for language growth (Newton, 2020). They will also provide the means to compare the effectiveness of song use with other more established materials for and methods of teaching and learning language, such as the use of dictionaries (Laufer & Hadar, 1997), graded readers (Horst et al., 1998; Warning & Takaki, 2003), academic lectures (Vidal, 2003; 2011), flashcards or words cards (Nakata, 2011), and television shows (Rodgers, 2013; Peters & Webb, 2018).

In response to this gap in empirical evidence surrounding song use for language development, there has been an emergence of research in this area in recent years. Alisaari and Heikkola (2016) studied how singing and listening to songs affect written fluency compared to reciting lyrics without music. They measured fluency based on the number of words used in the written task that was used as the pretest and posttest. They provided the lyrics or the written format of the sentences with the meanings to participants in all three conditions of their study. Their results showed that the students with the highest level of engagement with the song (i.e., productive use of the song by imitating aural input through singing with access to the lyrics) had the highest increase in fluency at 44%, followed by the other productive group (i.e., reciting lyrics with no music) at 32%, and finally the receptive only group (i.e., listening to the song) with the increase of 31% on the written fluency test. These results suggest that songs and lyric-based activities can promote fluency development. It is important to point out that while the non-musical group showed a slightly higher fluency increase between the pretest and posttest than the listening group, to some degree, this may be due to the congruency effect between the productive format of the practice condition and the test. Research has shown that productive practice leads to higher scores on a productive test than receptive practice (Lotto & de Groot, 1998; Schmitt, 2010; Webb, 2009).

In a large-scale (N=300) study with EFL Thai students, Pavia et al. (2019) investigated how listening to songs in a classroom setting affects incidental vocabulary learning. They found that participants showed increased learning of 10.97% for the spoken form of collocations and 6.53% for the spoken form of single-word items. They also suggested that the frequency of encounters through repeated listening significantly affected learning. These results provided

empirical evidence that songs may be an invaluable source of language input for acquiring multiword units (MWU) such as collocations.

Tomczak and Lew (2019) also looked at acquiring MWU through the use of songs in a language-focused activity in two studies. The participants in both studies deliberately practiced 13 multi-words and completed a pretest, an immediate posttest, and a one-week delayed posttest measuring productive recall. In the song groups, the target multi-words were embedded in a song, while the comparison groups learned the same multi-words in recorded spoken sentences. The results indicated the same learning pattern across the two studies, with the song groups outperforming the comparison groups on the immediate posttest by 5.33% and 4.57% and on the delayed posttest by 7.5% and 10.61%. Interestingly, the increased learning gap between the comparison and experimental groups on the delayed posttest results suggested that using songs may promote long-term retention of multi-word units compared to non-musical interventions.

Based on the available empirical evidence, using songs for language acquisition has been shown to have a positive effect. Nevertheless, the inconsistent patterns of use and selection of materials across studies and variance in the amount of potential learning from songs require further investigation in this field.

4.3.2 Importance of Attention Raising through Retrieval Activity for Vocabulary Learning

The plethora of research across cognitive science, psychology, education, and applied linguistics has led to the wide acceptance of the important role of attention in learning and retrieving new words (Treisman, 1969; Schmitt, 2008). Dating back to the 1950s, Broadbent's (1958) filter theory attempted to provide a unified theory of how selective attention operates and suggested that learners have limited capacity and only a limited portion of the stimuli pass the threshold between short-term memory and are stored in the long-term memory of individuals

based on the level of attention they receive. While this theory was not without flaws and did not account for the importance of quality and intelligibility of stimuli (Treisman, 1964; 1969), it highlighted the importance of attention in learning and prompted further research and the creation of frameworks in this area. In the 1970s, Craik and Lockhart (1972) argued that there was a need for a "new framework for research" (Lockhart & Craik, 1978. p.171) that accounts for how well new information is processed and proposed the *depths or levels of processing.* The concept of depth of processing "implies a greater degree of semantic or cognitive analysis." (Craik and Lockhart, 1972, p. 675); at this theory's core lies the quality of attention. Nevertheless, the ambiguous nature of this theory did not provide practical direction for testing nor clearly define what it meant to process new information deeply (Eysenck, 1978; Nelson, 1977). As Eysenck (1978) suggested, a new theory needed to account for the depth of processing in conjunction with the amount and the nature of the processing. As a result, in recent years, more operational theories such as the "Technique Feature Analysis" (TFA) (Nation & Webb, 2011) have emerged.

The *Technique Feature Analysis* (TFA) is another framework that provides practical evaluation for activities in vocabulary learning studies (Nation & Webb, 2011). The TFA framework consists of five main categories: motivation, noticing, retrieval, generation, and retention. These categories were then subdivided into 18 specific items. According to this framework, an activity that includes these specific items is more likely to promote learning. In a study, Zou et al. (2018) examined the predictability of TFA across four activities. They discovered that there's a positive relationship between the score of the activity on the TFA and the learning gains of the participants on the corresponding activity.

Retrieval, a key feature of the TFA, can be considered an effective attention-raising activity (Webb & Nation, 2017); retrieval activities can raise learners' attention to specific items that learners failed to retrieve from memory in an output task, thus giving learners an incentive to pay renewed attention to the (forgotten) information in the subsequent encounters. Further detail on the role of the sequence of input and output in activities is provided in the next section.

Moreover, text reconstruction activities can serve as a retrieval activity where learners are required to reproduce a text from memory (Boers, 2021). Izumi (2002) examined the effect of text reconstruction as an output task in an input-output-input cycle to help learners acquire a grammar pattern that repeatedly occurred in the input text. Their analysis indicated that the text reconstruction activity followed by revisiting the input text resulted in superior learning outcomes compared to input only activities.

In a more recent study, Yu et al. (2022) investigated learning multiword items through an input-output-input cycle using dictation, dictogloss, or answering comprehension questions as different types of output activities. Dictogloss, in which learners attempt to reconstruct a text from memory, showed the most promising learning outcomes. The dictation group also showed higher learning than the comprehension question activity group on the immediate posttest. However, these differences in learning gains decreased by the delayed posttest, and the dictation group no longer showed statistically significant learning gains compared to the comprehension question group (p=0.086). The authors attributed the better outcomes of the dictogloss activity relative to the dictation to the role of retrieval which is a key feature of the former activity.

The aforementioned theoretical frameworks and the empirical studies that followed them all make a compelling argument for the use of attention-raising techniques for increasing the quality of engagement, furthermore, output activities or combination of input and output

activities can increase the level of engagement and result in favourable learning outcomes. The current study aims to determine whether retrieval through text reconstruction can increase the quality of engagement in song-based activities and lead to vocabulary learning gains which have been the focus of criticism in song-based activities in the past.

4.3.3 Effects of Integration of Input and Output Activities on Vocabulary Learning

Krashen's (1981; 1985; 2004) comprehensible input hypothesis argues that learners learn what they hear and read through understanding the message. For the learning process to succeed, learners must encounter large quantities of comprehensible input in various contexts. This hypothesis relies heavily on the role of receptive knowledge in language development.

On the contrary, Swain (1985) argues that receptive knowledge is insufficient for deeper language learning levels that would allow for the productive use of language. Through her output hypothesis, she emphasized the critical role of productive language use/learning. She suggested that output can have three functions; one, through productive use, learners can evaluate their knowledge and notice the gaps. Two, through a trial-and-error process, they can produce language the best way possible, and based on the feedback received, they can modify their language knowledge. Three, by verbally discussing and reflecting on language features, learners can solve their language problems (Swain, 2005). Swain's hypotheses suggest that both receptive and productive language learning is needed to develop sufficient depth and breadth of language knowledge.

Nation (2007) combined the receptive and productive learning theories and created the *Four strands principle*. He suggested that a balance between meaning-focused input, meaning-focused output, language-focused learning, and fluency development strands, which

involve receptive and productive learning of all skills (reading, listening, writing, and speaking), are needed for substantial language learning.

The growing body of research investigating vocabulary learning supports this receptive and productive language learning integration. (Mondria &Wiersma, 2004; Newton & Nguyen, 2018; Nation, 2013; Webb, 2005; 2008; 2009). Zhang (2017) investigated the effects of receptive, productive, and receptive-productive tasks on collocations and grammar development. The results indicated that the combination of receptive-productive tasks lead to significantly ($p < .001$) higher learning outcomes.

In addition to the sheer presence of receptive and productive skills, or in other words, input and output, the sequence in which they occur may affect learning outcomes. Going back to Swain's (1985) output hypothesis, after the initial exposure (i.e., input), the push to produce the target language from memory (i.e., output) would allow the learners to identify the gaps in their knowledge and pay attention to the language if they were to be exposed to it again (i.e., input). The study by Nguyen and Boers (2019), investigating the effects of using the specific input-output-input sequence of activities on vocabulary learning from TED talks, confirmed that the use of retrieval, where the participants summarized what they had heard in the TED talk in a re-telling activity and then watched the TED talk a second time, resulted in significantly better vocabulary recall. Thus, to understand what can be learned from songs, it is important to investigate both receptive and productive language use through songs and how their integration and sequence of use could foster language learning.

4.4 Current Study

The present study investigates the effects of song use on promoting the learning of formulaic sequences (FS) in a classroom setting. Through two types of interventions, with and

without a retrieval activity in the specific sequence of input-input-output-output-input song use is explored. FS learning is operationalized as the participant's ability to recall the written productive form and recognize the target multi-word units' aural and written receptive form. All results are compared to those of the participants in the control group that did not receive any exposure to the target FS through any type of instruction. The control group participants completed the pretest, immediate posttest, and delayed posttest to control for any potential test effects.

4.5 Research Questions

4.5.1 Measuring Overall Vocabulary Learning

Does input-input-output-output-input sequential use of a song contribute to L2 formulaic sequences learning?

4.5.1.1 Measuring Receptive and Productive Knowledge Separately

What are the receptive and productive FS learning gains from the input-input-output-output-input sequential use of a song?

4.5.2 Effects of Retrieval on Vocabulary Learning

Which learning condition leads to higher learning gains in one session?

I. Listen, listen with lyrics, sing along with lyrics, sing along without lyrics, and finally, listen only.

II. Listen, listen with lyrics, sing along with lyrics, write down lyrics while listening to the instrumental version of the song, and finally, listen only.

4.6 Method

4.6.1 Participants

Sixty Thai EFL participants in grades 11 and 12 with ages ranging from 16 to 18 (*M*=17.66, *SD*=) were randomly assigned to one of three learning conditions in this study. Children in Thailand start learning English as early as three years old as a part of their preschool/ pre-elementary curriculum. Attending preschool is not compulsory; nevertheless, more than 74% of Thai children attend preschool across the country (Bureau of International Cooperation, 2008). The participants in this study had studied English for approximately 14 years (*M*=13.75, *SD*=) at their Thai schools (Public and Private). Their regular English teacher had placed them in the beginner-intermediate proficiency level. Nevertheless, during the first week of the study, all participants completed the updated Vocabulary Levels Test (VLT) (Webb et al., 2017) (See Appendix D) for a more accurate estimate of their vocabulary knowledge. The results indicated that the participants had knowledge of the first 1000 most frequent word families (*M* = 25.13 out of 30, *SD* = 3.08) and were learning the most frequent 2000 (*M* = 18.87 out of 30, *SD* = 4.63) and 3000 (*M* = 5.92 out of 30, *SD* = 3.75) word families. Previous research suggests that a correct response on the VLT represents knowledge of 33.3 words (Webb & Chang, 2015); thus, the participants in this study had knowledge of approximately 1,662.20 of the most frequent 3,000-word families.

This study used the pre-existing classes that the Thai school had randomly assigned each student to; therefore, the Kolmogorov-Smirnov normality test concerning participants' VLT scores was used to confirm the normal distribution of the participants across conditions. The results were insignificant ($p > 0.05$), confirming the normality of the distribution.

4.6.2 Research Instruments

This study used the research instruments from the study described in Chapter Three of this thesis. That includes the target song, FS items selected for testing, and the vocabulary knowledge test used for the pretest, immediate posttest, and delayed posttest. The participants in this study are approximately the same age, proficiency level, and number of years enrolled in English classes. Therefore, the same justification used in chapter three of this thesis for selecting the instruments was used in this study.

4.6.2.1 Song

The song selected in the previous chapter is the clean radio version of the song "Circles" by Post Malone (2019). The song and the lyrics were retrieved from YouTube https://www.youtube.com/watch?v=9gkRJONZPSA&list=RDGMEMHDXYb1_DDSgDsobPsOFxpAVM9gkRJONZPSA&index=1 (November 2020) (See Appendix H).

The lyrics of the song were analyzed by the software “Range” (Heatley & Nation, 2002) and Nation's (2017) British National Corpus/ Corpus of Contemporary American English (BNC/COCA) word family lists" (p.7). Table 4.1 provides the lexical profile of the target song. The analysis suggests that target students in this study, who knew approximately 1662 of the most frequent word families, would have lexical coverage of 93.09%-98.80% of the target song. According to previous research on the relationship between lexical coverage and incidental learning without assistance, 95% lexical coverage is needed for learning to occur from auditory sources of language input (van Zeeland & Schmitt, 2013). Therefore, the selected song is at the appropriate difficulty level for the participants to learn the target FS incidentally. For further details on the justification for selecting the target song, please refer to section 3.6.2.1 in this volume.

Table 4.1 Lexical frequency Profile: “Circles.”

Level	Token%	Cumulative coverage %
1	93.09	93.09
2	5.71	98.80
3	0.60	99.40
4	0.00	99.40
5	0.60	100.00

Similar information is found in Table 3.1 of this volume (p.74)

4.6.2.2 Target Sequences (FS)

The selected 16 FS are presented in Table 4.2. For further details on the justification for selecting the target FS, please refer to section 3.6.2.2 in this volume.

Table 4.2 Target FS

Song	Target FS	Number of encounters in the song	*fr*	MI Score
Circles	Turn Around	1	28591	4.61
	Upside Down	1	5843	7.40
	Feed the Flame	2	86	2.54
	Every Time	1	38792	-
	Seasons Change	2	104	3.73
	Let Go/ Let it Go	4	14028	5.81
	Run Away	6	7031	8.49
	Running in Circles	2	233	3.50
	Take the Blame	2	1154	2.73
	Got a Feeling	2	605	-

From the Get-go	1	996	-
I'm going Through	1	19965	2.54
Waiting On	2	3295	-
Don't understand	1	3574	-
I said so	1	12612	-
Make up your mind	1	654	-

Similar information is found in Table 3.2 of this volume (p.75)

4.6.2.3 Dependent Measures

The vocabulary knowledge described in section 3.6.2.3 of this volume was used for all groups during the pretest, immediate posttest, and two-week delayed posttest (see Appendix I).

4.6.3 Procedure

The three student groups were assigned to either control (C) or one of the experimental groups (E1 & E2). In Chapter One, Section 1.3.4, the procedure for Week One was outlined. This included collecting consent forms, demographic information forms, VLT, and the vocabulary knowledge test for the pretest.

During week two, the control group (C) completed the immediate posttest without any exposure to the target song. The experimental groups (E1 & E2) participated in one of the two sequential learning conditions. Participants in the E1 group listened, listened again while reading the printed lyrics, sang along with the lyrics, sang along with the song without access to the printed lyrics, and listened again. Participants in the E2 group listened, listened again while reading the printed lyrics, sang along with the lyrics, wrote down the lyrics while listening to the instrumental version of the song from memory, and finally listened to the song one last time. The participants' text-reconstruction work from the last output portion of the activity was collected

prior to their last listening session. This ensured that the participants could not make any changes to the text they had previously produced after listening to the target song one last time. During this session, participants completed the immediate posttest following the learning conditions. It is important to note that participants were asked to complete all tasks individually. They were not prewarned of the structure or nature of the different activities. For example, the participants in E2 did not know they would be asked to reconstruct the lyrics prior to performing this task. This ensured that the level of attention to the lyrics during the initial rounds of input were not influenced by their knowledge of the follow-up activities. Also, in the event that the participants asked questions that could have influenced the results of the study, they were told that all of their questions would be answered at the end of our last session together. This included queries about the name of the song or about the meaning of a particular word in the song.

During week three, all participants went back to their regular Thai school curriculum with their Thai teachers. Finally, on week four, all control and experimental groups completed the two-week delayed post-test. Table 4.3 provides an overview of the procedure.

Table 4.3 Procedure Timeline

Learning Condition	Week 1	Week 2	Week 3	Week 4
• Control (C)	• Consent Form • Demographic • information form • VLT	• Immediate Posttest	• Regular Thai school curriculum	• Delayed Posttest
• Experimental 1 (E1)		• Listen, Listen with lyrics, Sing along with lyrics, Sing along without lyrics, and finally Listen only • Immediate Posttest	• Regular Thai school curriculum	• Delayed Posttest

• Experimental 2 (E2)		• Listen, Listen with lyrics, Sing along with lyrics, Write down lyrics while listening to the instrumental version of the song, and finally, Listen only • Immediate Posttest	• Regular Thai school curriculum	• Delayed Posttest

4.6.4 Analysis

IBM SPSS Statistics version 28 data analysis was used to generate the descriptive statistics, including means, standard deviation, etc., and inferential statistics, such as the multivariate analysis of variance (MANOVA), to compare the results between pretest, immediate posttest, and two weeks delayed posttest within and between groups. To assume homogeneity of within-group variance, Mauchly's Sphericity test was used. If the assumption of sphericity had been violated, the Greenhouse-Geisser was used to adjust the *df* (Warner, 2013).

4.7 Results

Tables 4.4 and 4.5 provide descriptive statistics based on the participant's scores on the vocabulary knowledge test on the pretest, immediate posttest, and the two weeks delayed posttest. Table 4.4 illustrates the overall knowledge of FS while table 4.5 separates receptive and productive aspects of FS knowledge. Each receptive knowledge and productive knowledge test had a possible maximum score of 16, resulting in a total of 32 for overall knowledge of formulaic sequences (FS). To investigate the impact of exposure to a song on the receptive, productive, and overall learning of FS, the first two research questions were answered using repeated measures ANOVA. The analysis considered the interaction between time and the sequence of input-input-output-output-input activities

Table 4.4 Descriptive Statistics of Overall Vocabulary Test Scores: Means (SDs)

Participant subgroups	Total FS Knowledge		
	Pretest	Immediate posttest	Delayed posttest
C (*n= 17)*	2.97(1.69)	3.23(1.60)	2.82(1.24)
E1 (*n= 19)*	2.68(2.00)	18.63(4.03)	15.47(4.30)
E2 (*n = 24)*	3.54(2.24)	24.04(5.64)	20.33(5.28)

Note. The maximum score on both the productive and receptive tests was 32.
Note. C= Control, E1= Listen, listen with lyrics, sing along with lyrics, sing along without lyrics, and finally, listen only, E2= Listen, listen with lyrics, sing along with lyrics, write down lyrics while listening to the instrumental version of the song, and finally, listen only.

Table 4.5 Descriptive Statistics of Receptive and Productive Vocabulary Test Scores: Means (SDs)

Participant subgroups	Productive FS Knowledge			Receptive FS Knowledge		
	Pretest	Immediate posttest	Delayed posttest	Pretest	Immediate posttest	Delayed posttest
C (*n= 17)*	1.09(0.83)	1.00 (0.86)	0.76 (0.90)	1.88(1.05)	2.23(1.09)	2.06(0.83)
E1 (*n= 19)*	0.79(1.27)	6.37(3.18)	5.00(3.15)	1.89(1.33)	12.26(1.75)	10.47(2.14)
E2 (*n = 24)*	1.17(1.47)	10.37(3.56)	11.13(2.72)	2.37(1.17)	13.67(2.66)	11.13(2.72)

Note. The maximum score on the productive test and the receptive test was 16.
Note. C= Control, E1= Listen, listen with lyrics, sing along with lyrics, sing along without lyrics, and finally, listen only, E2= Listen, listen with lyrics, sing along with lyrics, write down lyrics while listening to the instrumental version of the song, and finally, listen only.

4.7.1 Overall Vocabulary Learning from input-input-output-output-input sequential use of songs

According to the analysis, the control group did not show any learning between the pretest, posttest, and delayed posttest. However, the experimental groups showed improvement in their overall learning between these tests, as illustrated in Figure 4.1. The analysis for the control group showed that the assumption of sphericity had been met $\chi^2(2) = 3.302$, $p = 0.192$, and the within-participant main effect (time) was not statistically significant $F(2, 32) = 1.224$, $p = 0.307$.

The analysis for the non-retrieval group (E1) showed that the assumption of sphericity had been violated, $\chi^2(2) = 7.77$, $p = 0.021$; therefore, a Greenhouse-Geisser was used. The results indicated that the within-participant main effect (time) was statistically significant, $F(1.46, 26.33) = 358.094$, $p < 0.001$ with partial $\eta^2 = 0.952$ (Very large effect size). As for the retrieval group (E2), the assumption of sphericity had been met $\chi^2(2) = 7.898$, $p = 0.019$, and the within-participant main effect (time) was statistically significant, $F(2, 46) = 252.22$, $p < 0.001$ with partial $\eta^2 = 0.916$ (Very large effect size).

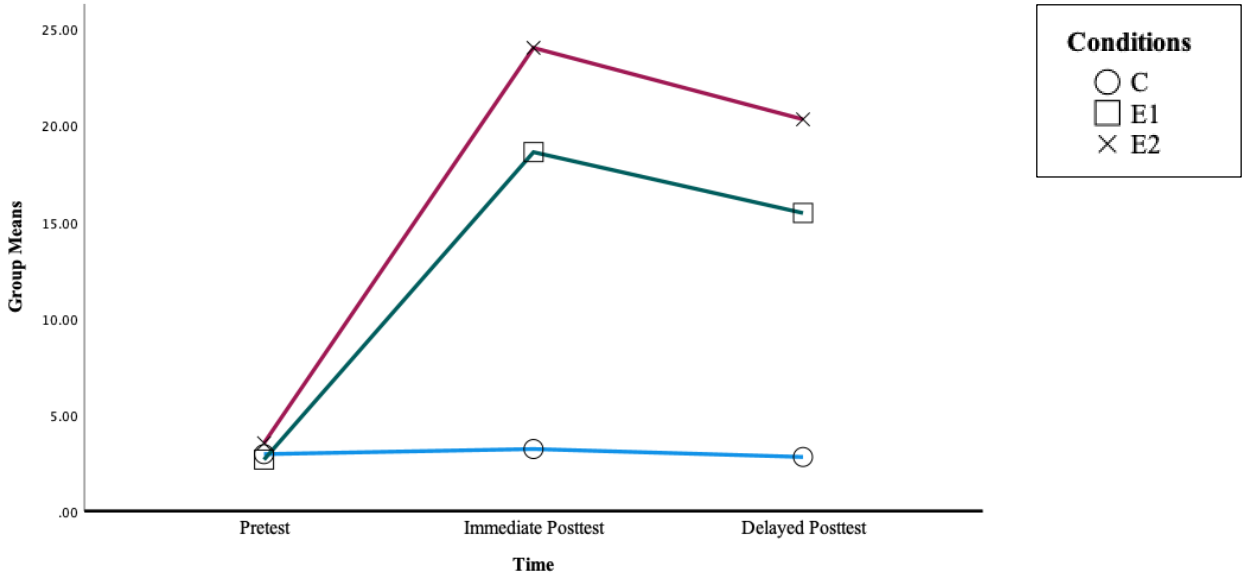


Figure 4.1 Group means for overall vocabulary learning over time.

4.7.2 Measuring Receptive and Productive Knowledge

The analysis found that there were no learning gains for group C, but both groups E1 and E2 showed improvements in their scores on both the Receptive and Productive tests, as seen in Figures 4.2 and 4.3. For group C, the analysis of the scores on the receptive test indicated that the assumption of sphericity had been met $\chi^2(2) = 3.005$, $p = 0.223$, and the within-participant main effect (time) was not statistically significant $F(2, 32) = 2.227$, $p = 0.124$. For the productive test, the analysis showed that the assumption of sphericity had been violated, $\chi^2(2) = 13.227$, $p = 0.001$, and therefore, a Greenhouse-Geisser was used. The results indicated that the within-participant main effect (time) was not statistically significant $F(1.261, 20.17) = 1.93$, $p = 0.179$.

On the receptive test, the analysis for E1 indicated that the assumption of sphericity had been violated, $\chi^2(2) = 8.203$, $p = 0.017$, and therefore, a Greenhouse-Geisser was used. The results indicated that the within-participant main effect (time) was statistically significant, $F(1.446, 26.034) = 491.63$, $p < 0.001$ with partial $\eta^2 = 0.965$ (Very large effect size). For E2, the analysis revealed that the assumption of sphericity had been met $\chi^2(2) = 0.723$, $p = 0.697$, and the

within-participant main effect (time) was statistically significant, $F(2, 46) = 252.950$, $p < 0.001$ with partial $\eta^2 = 0.917$ (Very large effect size).

As for the productive test, the analysis for E1 indicated that the assumption of sphericity had been violated, $\chi^2(2) = 15.057$, $p < 0.001$; therefore, a Greenhouse-Geisser was used. The results showed that the within-participant main effect (time) was statistically significant, $F(1.26, 22.67) = 61.04$, $p < 0.001$ with partial $\eta^2 = 0.772$ (Very large effect size). For E2, the analysis revealed that the assumption of sphericity had been violated $\chi^2(2) = 19.042$, $p < 0.001$, and therefore, a Greenhouse-Geisser was used. The results indicated that the within-participant main effect (time) was statistically significant, $F(1.266, 29.129) = 137.327$, $p < 0.001$ with partial $\eta^2 = 0.857$ (Very large effect size).

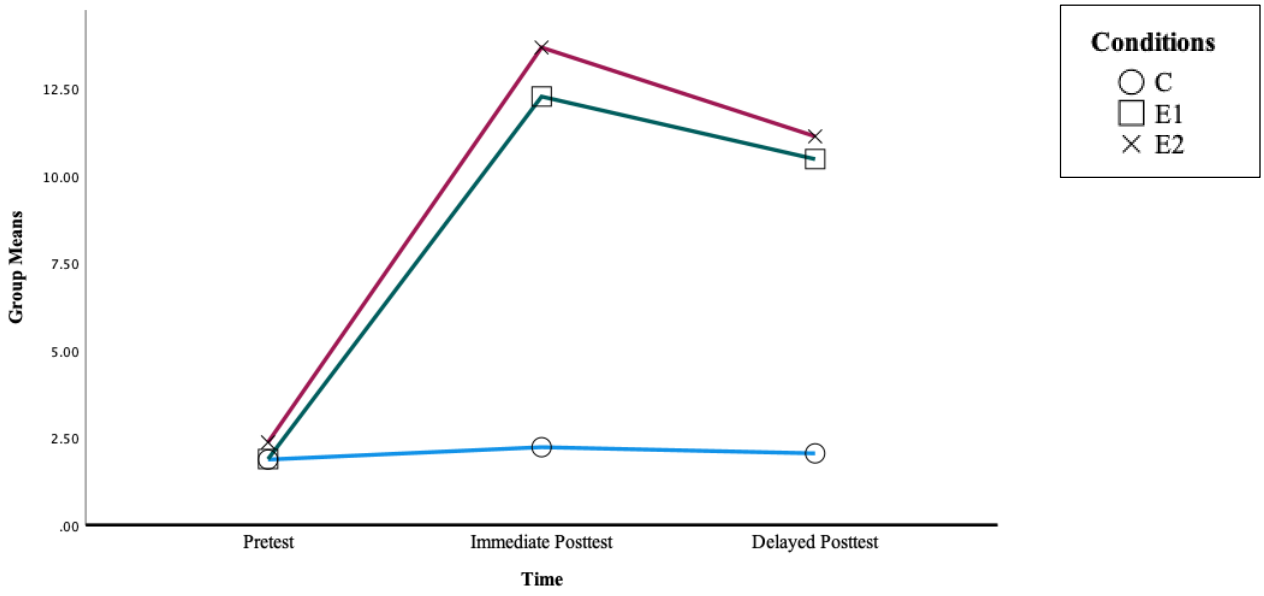


Figure 4.2 Group means for receptive vocabulary learning over time.

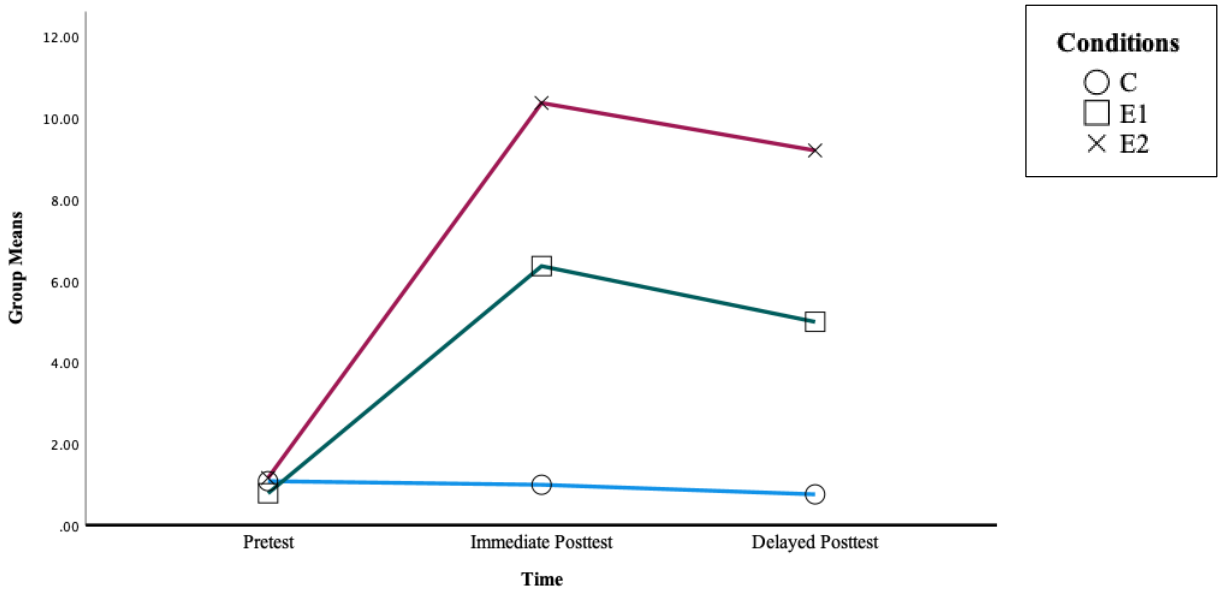


Figure 4.3 Group means for productive vocabulary learning over time.

4.7.3 FS Learning Gains

A series of pairwise comparisons were run to further explore the differences in the receptive, productive, and total FS learning between groups at different times of testing (Pretest, immediate posttest, and two weeks delayed posttest). Consistent with the results in the previous section, no significant differences were found for the control group, while the experimental groups showed significant gains with extremely large (d >2.0) effect sizes. See further details on the pairwise comparison scores of the pretest, immediate posttest, and delayed posttest in Table 4.6.

Table 4.6 Pairwise comparison for different sections of the tests

Time of testing (*i*)	Time of testing(*j*)	Difference between means (*j*-*i*)	*SD* error	*p*	Cohen's *d*	95% confidence interval for the difference	
						Upper bound	Lower bound
Control Receptive							
1	2	0.353	0.191	0.248	0.24	0.862	-0.157
	3	0.176	0.176	0.997	0.19	0.648	-0.295
Control Productive							
1	2	-0.088	0.107	1.00	0.11	0.198	-0.374
	3	-0.324	0.223	0.496	0.38	0.271	-0.918
Control Total							
1	2	0.265	0.243	0.876	0.16	0.914	-0.385
	3	-0.147	0.320	1.00	0.10	0.709	-1.003
E1 Receptive							
1	2	10.368*	0.317	<0.001	6.67	11.206	9.531
	3	8.579*	0.448	<0.001	4.82	9.761	7.397
E1 Productive							
1	2	5.579*	0.632	<0.001	2.30	7.247	3.911
	3	4.211*	0.605	<0.001	1.75	5.807	2.614
E1 Total							
1	2	15.947*	0.664	<0.001	5.01	17.701	14.194
	3	12.789*	0.763	<0.001	3.81	14.803	10.776
E2 Receptive							
1	2	11.292*	0.569	<0.001	5.49	12.762	9.822
	3	8.750*	0.519	<0.001	4.18	10.089	7.411
E2 Productive							
1	2	9.208*	0.761	<0.001	3.39	11.173	7.244
	3	8.042*	0.645	<0.001	4.56	9.706	6.377
E2 Total							
1	2	20.50*	1.178	<0.001	4.77	23.541	17.459
	3	16.792*	0.982	<0.001	4.14	19.328	14.256

Note. C= Control, E1= Listen, listen with lyrics, sing along with lyrics, sing along without lyrics, and finally, listen only, E2= Listen, listen with lyrics, sing along with lyrics, write down lyrics while listening to the instrumental version of the song, and finally, listen only.

4.7.4 The Effects of Retrieval on FS Learning

In response to the final research questions, investigating the effects of retrieval on learning FS while using songs, Multivariate analysis of variance (MANOVA) was used.

4.7.4.1 Immediate Posttest

The results indicated an overall statistically significant difference, F (4, 114) = 32.514, $p < 0.001$, Pillais' Trace = 1.066, partial $\eta^2 = 0.533$ on the immediate posttest. The analysis for different sections of the test considered separately showed significant differences across sections. Both receptive recognition (F (2, 57) = 172.97; $p < 0.001$; partial $\eta^2 = 0.859$) and productive recall (F (2, 57) = 51.295; $p < 0.001$; partial $\eta^2 = 0.643$) showed extremely large effect sizes.

To determine whether having the retrieval activity influenced the performance on different sections of the test, the Post hoc comparison using the Bonferroni test was used. Since no learning occurred at all in the control group, this section zooms in on the differences between the two experimental groups, The retrieval group (E2) outperformed the non-retrieval group (E1) significantly on the productive test ($p < 0.001$, d = 1.19) with a very large effect size. However, no significant differences were found between the retrieval and non-retrieval groups on the receptive test.

4.7.4.2 Delayed Posttest

The results indicated an overall statistically significant difference, F (4, 114) = 33.984, $p < 0.001$, Pillais' Trace = 1.088, partial $\eta^2 = 0.544$ on the delayed posttest. The analysis for different test sections considered separately showed significant differences across sections. Receptive recognition (F (2, 57) = 101.722; $p < 0.001$; partial $\eta^2 = 0.781$) and productive recall (F (2, 57) = 48.466; $p < 0.001$; partial $\eta^2 = 0.630$), both showed extremely large effect sizes.

To determine whether having the retrieval activity influenced the performance on different sections of the test, the Post hoc comparison using the Bonferroni test was used. Same trends as the results on the immediate posttest were observed. The retrieval group (E2) outperformed the non-retrieval group (E1) significantly on the productive test ($p < 0.001$, d =

2.08) with an extremely large effect size. However, no significant differences were found between the retrieval and non-retrieval groups on the receptive test.

4.8 Discussion

4.8.1 Does the input and output sequence integration contribute to FS learning?

In response to the first research question investigating the incidental learning of formulaic sequences from a song through a particular input-input-output-output-input sequence, the results showed that repeated exposure (i.e., Five times) to a song in this sequence of input, output, and input again, has the potential to foster FS learning. The participants in the non-retrieval (E1) group showed learning gains of 49.84% (15.95 items), and the retrieval (E2) group showed learning gains of 64.06 (20.50 items) from the pretest to the immediate posttest. They also showed retention of 39.97% (12.79 items) and 52.47 (16.79 items) on the two weeks delayed posttest. In contrast, participants in the control group did not show any statistically significant learning gains, with gains of 0.81% (0.26 items) and -0.47% (-0.15 items) on the immediate and two weeks delayed posttests, respectively.

Measuring the potential learning of receptive and productive knowledge of FS separately from sequential use of input and output of songs, found similar learning patterns between different study conditions compared to the overall knowledge of FS.

The results from the receptive recognition test between the pretest and the immediate post revealed significant learning gains for the experimental groups, 64.81% (10.34 items) for E1 and 70.63% (11.3 items) for E2. These participants retained a large portion of their learning gains by the two weeks delayed posttest, with learning gains of 53.63% (8.68 items) for E1 and 54.75% (8.76 items) for E2. In contrast, the control group's minuscule learning gains of 2.19% (0.35

items) from the pretest to the immediate posttest and 1.13% (0.18 items) from the pretest to the delayed posttest were statistically insignificant.

The results from the productive recall test between the pretest and the immediate post revealed significant learning gains for the experimental groups, 34.88% (5.58 items) for E1 and 57.50% (9.2 items) for E2. These participants retained a large portion of their learning gains by the two weeks delayed posttest, with learning gains of 26.31% (4.21 items) for E1 and 50.26% (8.42 items) for E2. In contrast, the control again did not show any significant learning gains.

The current study's results are consistent with those found in the previous two studies in this thesis as well as other studies in the literature (Coyle & Gómez Gracia, 2014; Medina, 1993; Pavia et al., 2019; Tegge, 2015; Tomczak and Lew, 2019), suggesting that songs have the potential to promote learning of vocabulary items and, more specifically, FS. However, the learning gains in the current study seem a bit higher than those found in the previous two studies of this volume. Considering that the target participants in the three studies of this volume were similar with regard to their VLT scores, L1, country of origin, age group, and attended the same school, it is reasonable to assume that the differences found between studies stem from differences in treatment conditions, Study one of this thesis investigated the effects of repeated listening (i.e., Five times) to a target song in mass and spaced listening sessions. The results revealed learning gains of 34.57% (2.35 collocations) and 34.42% (2.412 collocation) on the immediate posttest and showed retention of 28.57% (2.00 collocations) and 32.71% (2.94 collocations), respectively, on the two weeks delayed posttest of the collocation recognition test. These results show considerably lower learning gains than those found in the current study, as indicated by the receptive recognition test scores. This variance may be due to several reasons. First, it is important to note that the number of items used for testing collocation recognition was

less than half (i.e., Seven collocations tested) compared to the number of FS items used in this study (i.e., 16 items tested for the receptive recognition of FS). Second, study one only had the participants listen to the target song without access to the lyrics or a push for production. As illustrated in the earlier section of this article, based on the output hypothesis proposed by Swain (1985; 2005), the output can play an invaluable role in promoting learning; It is through the combination of input and output that learners realize the gaps in their knowledge and are able to fill those gaps in the subsequent exposure. A study conducted by Alisaari and Heikkola (2016) investigated the impact of songs and lyric-based instruction on writing fluency; Their results revealed a 44% improvement in the singing condition and a 31% improvement in the listening-only condition illustrating the importance of output practice in encouraging learning. In addition, the results from study two of this volume provided further support for the role of output, considering that the only statistically significant difference between the experimental groups on the FS recognition test was between the group that listened while reading the lyrics (LL) and the group that listened and sang while reading the lyrics (LSL), with the LSL group outperforming LL group by 13.8% (2.22 items).

Study two of this thesis which provided the instruments for the current study, investigated learning the receptive and productive aspects of FS from different modes of input using a song. The results indicated that listening and singing while reading the lyrics (LSL) five times provided the best learning outcomes, with 58.75% (9.4 items) learning gains on the immediate posttest for form recognition and 56.56% (9.05 items) on the immediate posttest for form recall. Furthermore, the participants retained a large portion of their learning gains on the two weeks delayed posttest, with 56.88% (9.1 items) on the form recognition test and 54.06% (8.65 items) on the form recall test. The LSL condition in study two is the closest condition that can be

compared to the E1 condition in the current study, with the difference that after the first exposure, which only included listening to the song, the subsequent exposures included all the listening, reading, and signing modes at the same time. Also, study two participants were exposed to the target song in spacing intervals. In comparison, the current study used a sequential format where the output was only present in two out of four rounds of exposures, and they were followed by another round of input at the end. Furthermore, the current study exposed the participants to the target song in one session, followed by the immediate posttest. Nevertheless, the learning gains from these conditions (i.e., LSL in study two and E1 in the current study) are extremely close to the receptive aspect of word knowledge. E1 showed slightly higher learning gains of 6.06% (0.97 items) on the immediate posttest, while LSL showed slightly higher gains of 3.25% (0.52 items) on the two weeks delayed posttest. However, the form-recall test scores showed that the LSL group did considerably better, with higher gains of 21.68% (3.47 items) on the immediate posttest and 27.75% (4.44 items) on the two weeks delayed posttest.

The similar outcomes between the LSL in study two and E1 in the current study for the receptive aspect of FS knowledge, as indicated by the scores on the form recognition test, can be explained by the congruency theory between learning conditions and learning outcomes. To this point, both learning conditions included the same format and number of exposures (i.e., Five times) for receptive learning, meaning that they had auditory exposure through listening, visual exposure through reading the lyrics, or both, at every exposure; therefore, it is not surprising that these similar receptive learning conditions in both groups led to similar receptive learning of the target FS. In a study by Webb (2009) investigating the effects of receptive and productive learning of word pairs on learning different receptive and productive aspects of target words, he found that receptive learning conditions lead to the receptive aspect of word knowledge, while

productive learning can lead to the productive and receptive aspect of word knowledge. This may also explain why the LSL group scored higher on the productive recall test than E1. Considering that the LSL group had productive learning through singing in four exposures while E1 only had two, the differences between the productive learning conditions may have led to different learning outcomes.

The sequential order of input and output for E1 that did not include a retrieval condition did not seem to add any value when compared to the LSL group in the previous study. This may suggest that in learning conditions where the sequence of exposure does not enhance the quality of engagement, the frequency of exposure may have a more positive effect on learning.

4.8.2 Does adding a retrieval activity to the input, output, and input again sequences of exposure enhance FS learning?

In response to the final research question investigating the effects of retrieval on FS learning through sequential exposure to a song, the results showed statistically significant differences between the non-retrieval (E1) and retrieval (E2) groups on the productive aspect of word knowledge with E2 outperforming E1 by 25.04% (4.01 items) on the immediate posttest and 26.3% (4.21 items) on the two weeks delayed posttest. However, no statistically significant differences were found between the groups on the receptive aspect of word knowledge.

These results are consistent with previous studies. In chapter three of this volume, only one statistically significant difference was found between learning conditions on the receptive aspect of word knowledge. This lack of variance which may have been due to the ceiling effect, may also be the reason behind the statistically insignificant difference between the E1 and E2 conditions in this study. The receptive learning conditions, even in the least engaging group, may have been sufficient for learning the receptive knowledge of the target FS as measured by the

form recognition test. However, the statistically significant difference between E1 and E2 on the productive aspect of word knowledge, which is more difficult to acquire and involves deeper processing of information (González-Fernández & Schmitt, 2020; Hulstijn & Laufer, 2001; Webb, 2009), suggests that the added retrieval activity in the E2 group enhances the learning of FS from songs.

Going back to the discussion of increasing the quality of engagement and whether higher quality can increase learning gains with less exposure, the E2 group in this study can be compared to the LSL group in study two. As suggested in the last section, E1 showed lower learning gains for the productive aspect of FS than the LSL group, which may have been due to the higher frequency of exposure to the productive aspect of word knowledge. However, the E2 group, which had the same number of exposures as E1 to the productive aspect of word knowledge, showed approximately the same learning gains as LSL. E2 scored 0.94% (0.15 items) and 8.19% (1.31 items) higher than LSL on the immediate and delayed posttests. While LSL had an additional two exposures than the E1 and E2 groups in this study, the retrieval activity in E2 may have increased the quality of engagement in E2, leading to similar results to LSL condition with fewer exposures.

4.9 Conclusion and Future Direction

These findings provide further evidence that increased involvement with the target language through retrieval may increase the quality of engagement when using songs for incidental vocabulary learning. However, this study is not without limitations. First, the participants in this study were exposed to the target song in a mass practice session where there were no intervening time or unrelated activities. Research suggests that spaced practice, with a period between practice sessions and retrieval sessions, may result in more durable learning

(Cepeda et al., 2009; Kim & Webb, 2022). To this point, the comparisons made between the results of this study and those found in study two in the previous section should be interpreted cautiously, given the difference in the distribution of practice between the two studies. Considering the recommendations from research on distributed practice, it would be expected that spaced practice, including the format of the LSL group in study two, would show higher learning than the experimental groups in the current study. Since this was not the case, the results of this study, especially the outcomes of the E2 group, further emphasize the positive effect retrieval practice can have on language acquisition. Nevertheless, more research is needed to explore the effects of retrieval activities combined with spaced practice on song-based vocabulary teaching and learning.

Another area of research that needs further attention is using songs to promote listening comprehension and learning the form-meaning connection aspect of FS and single-word items. It can be argued that knowledge of the aural form of the single words and FS can potentially "prime" the learner to recognize these units when encountering them in other possibly more informative contexts. This priming effect can potentially give the learners a better chance of establishing the meanings at these subsequent encounters. The first study's results in this volume showed evidence that learning the form-meaning connection aspect of single words from songs is possible after repeated exposures considering that the learners had reached the appropriate lexical coverage of the context. However, the learning gains were small and further research is needed to explore other factors that may encourage the learning of form-meaning connections.

Furthermore, songs use simple and conversational language at approximately half the speed of spoken discourse with 75.49 WPM (Murphey, 1992b). In addition, the inherent intertextual nature of songs as a form of pop culture allows learners to bring their own life

experiences to the activity and connect with the lyrics instantly. For decades, this has led teachers and educators to use songs and lyric-based activities for language development in L1 language teaching (Johnson & Goering, 2016). Using songs provides an opportunity to experience the target language's culture and motivates L2 learners to engage with the lyrics and understand the story behind them. According to Nation (2007), one important factor for successful language acquisition through listening and reading is the learner's intrinsic motivation, which comes from their interest in the language input source and their desire to understand it. In sum, further research is required to discover empirically based methodologies that effectively focus on utilizing the narrative nature of songs that can potentially foster comprehension and learning of the form-meaning connection aspect of FS and single-word units.

4.10 References

Alisaari, J., & Heikkola, L. M. (2016). Increasing fluency in L2 writing with singing. Studies in Second Language Learning and Teaching, 6(2), 271–292. https://doi.org/10.14746/ssllt.2016.6.2.5

Arnold, J. L., & Herrick, E. (2017). *New ways in teaching with music.* TESOL International Association. Alexandria, Virginia, USA: TESOL press.

Broadbent, D. E. (1958). *Perception and communication.* Pergamon Press

Boers, F. (2020). Factors affecting the learning of multiword items. In S. Webb (Ed.). *The Routledge Handbook of vocabulary studies.* Routledge.

Boers, F. (2021). *Evaluating second language vocabulary and grammar instruction: A synthesis of the research on teaching words, phrases, and patterns*. Routledge

Bureau of International Cooperation. (2008). Towards a learning society in Thailand: An introduction to education in Thailand. *Ministry of Education.* Retrieved from https://www.bic.moe.go.th/images/stories/book/ed-eng-series/intro-ed08.pdf

Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., & Pashler, H. (2009). Optimizing distributed practice: theoretical analysis and practical implications. *Experimental Psychology, 56*(4), 236–246. https://doi.org/10.1027/1618-3169.56.4.236

Coyle, & Gómez Gracia, R. (2014). Using Songs to Enhance L2 Vocabulary Acquisition in Preschool Children. *ELT Journal, 68*(3), 276–285. https://doi.org/10.1093/elt/ccu015

Craik, & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

de la Fuente, M. J., (2002). Negotiation and acquisition of L2 vocabulary. *Studies in Second Language Acquisition, 24*(1), 81–112. https://doi.org/10.1017/S0272263102001043

Engh, D. (2013). Why use music in English language learning? A survey of the literature. *English Language Teaching, 6*(2), 113–127. doi:10.5539/elt.v6n2p113

Eysenck, M. W. (1978). Levels of processing: A critique. *69(2), 157–169.* *https://doi.org/10.1111/j.2044-8295.1978.tb01643.x*

González Fernández, B. (2022). Vocabulary knowledge: Lexical depth and its relationship with out-of-class exposure. In P. Szudarski, & S. Barclay (Eds.), *Vocabulary Theory, Patterning and Teaching* (pp. 21-47). SLA

González Fernández, B., & Schmitt, N. (2015). How much collocation knowledge do L2 learners have?: the effects of frequency and amount of exposure. ITL, 166(1), https://doi.org/10.1075/itl.166.1.03fer

González-Fernández, B., & Schmitt, N. (2020). Word knowledge: Exploring the relationships and order of acquisition of vocabulary knowledge components. *Applied Linguistics, 41*(4), 481–505. https://doi.org/10.1093/applin/amy057

Good, A. J., Russo, F. A., & Sullivan, J. (2015). The efficacy of singing in foreign-language learning. *Psychology of Music, 43*(5), 627–640. https://doi.org/10.1177/0305735614528833

Heatley, A., Nation, P., & Coxhead, A. (2002). Range and frequency programs. Available at: http://www.victoria.ac.nz/lals/staff/paul-nation.aspx.

Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *The Canadian Modern Language Review/Revue Canadienne des Langues Vivantes, 61*, 355–382. Doi: https://doi.org/10.1353/cml.2005.0018.

Horst, M., Cobb, T., & Meara, P. (1998). Beyond a Clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, *11*(2), 207–223.

Huang, S., Willson, V., & Eslami, Z. (2012). The Effects of Task Involvement Load on L2 Incidental Vocabulary Learning: A Meta-Analytic Study: The Effects of Task Involvement Load on L2 Incidental Vocabulary Learning. The Modern Language Journal (Boulder, Colo.), 96(4), 544–557. https://doi.org/10.1111/j.1540-4781.2012.01394.x

Hulstijn, J. (2001). Intentional and incidental second-language vocabulary learning: reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), Cognition and second language instructions (pp. 258–286). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9781139524780.011

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning, 51,* 539–558. https://doi.org/10.1111/0023-8333.00164

Izumi, H. (2002). Output, input enhancement, and the noticing hypothesis: An Experimental Study on ESL Relativization. Studies in Second Language Acquisition, 24(4), 541–577. https://doi.org/10.1017/S0272263102004023

Jin, Z., & Webb, S. (2020). Incidental vocabulary learning from listening to teacher talk. *Modern Language Journal*, *104*(3), 550-566. https://doi.org/10.1111/modl.12661

Johnson, L. L., & Goering, C. Z. (2016). Recontextualized [electronic resource] : A Framework for Teaching English with Music (Johnson & C. Z. Goering, Eds.; 1st ed. 2016.). SensePublishers. https://doi.org/10.1007/978-94-6300-606-4

Kim, S., & Webb, S. (2022). The Effects of Spaced Practice on Second Language Learning: A Meta-Analysis. *Language Learning*, *72*(1), 269–319. https://doi.org/10.1111/lang.12479

Krashen, S. D. (1981). The case for narrow reading. *TESOL News, 12*, 23.

Krashen, S. D. (1985). Input hypothesis: Issues and implications. London: Longman.

Krashen, S. D. (2004). The Case for Narrow Reading. Language (Los Angeles, Calif.), 3(5), 16-19.

Lai, C., Zhu, W., & Gong, G. (2015). Understanding the quality of out- of- class English learning. *TESOL Quarterly, 49*(2), 278– 308. https://doi.org/10.1002/tesq.171

Laufer, B., & Hadar, L. (1997). Assessing the effectiveness of monolingual, bilingual, and "bilingualised" dictionaries in the comprehension and production of new words. *The Modern Language Journal*, *81*(2), 189–196. https://doi.org/10.1111/j.1540-4781.1997.tb01174.x

Laufer, B., & Hulstijn, J. H. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*, 1–26. https://doi.org/10.1093/applin/22.1.1

Lindgren, E., & Muñoz, C. (2013). The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism,10*, 105–129. https://doi.org/10.1080/14790718.2012.679275

Lockhart, R. S., & Craik, F. I. M. (1978). Levels of processing: A reply to Eysenck. *The British Journal of Psychology, 69*(2), 171–175. https://doi.org/10.1111/j.2044-8295.1978.tb01644.x

Lotto, L., & de Groot, A. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning, 48*(1), 31–69. https://doi.org/10.1111/1467-9922.00032

Mondria, J.-A., & Wiersma, B. (2004). Receptive, productive and receptive + productive L2 vocabulary learning: What difference does it make? In P. Bogaards & B. Laufer (Eds.),

Vocabulary in a second language: Selection, acquisition, and testing (pp. 79–100). John Benjamins.

Murphey, T. (1992a). *Music and song.* In A. Maley (Ed.), Resource books for teachers. Oxford: Oxford University Press**.**

Murphey, T. (1992b). The discourse of pop songs. *TESOL Quarterly, 26*(4), 770-774. doi:10.2307/3586887

Nakata. (2011). Computer-assisted second language vocabulary learning in a paired-associate paradigm: a critical investigation of flashcard software. Computer Assisted Language Learning, 24(1), 17–38. https://doi.org/10.1080/09588221.2010.520675

Nakata, T. (2017). Does repeated practice make perfect? The effects of within-session repeated retrieval on second language vocabulary learning. *Studies in Second Language Acquisition, 39*, 653–679. https://doi.org/10.1017/S0272263116000280

Nakata, T., Tada, S., Mclean, S., & Kim, Y. A. (2020). Effects of distributed retrieval practice over a semester: Cumulative tests as a way to facilitate second language vocabulary learning. *TESOL Quarterly*, *55*(1), 248–270. https://doi.org/10.1002/tesq.596

Nation, I. S. P. (2007). The four strands. *International Journal of Innovation in Language Learning and Teaching, 1,* 1–12.

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). New York, NY: Cambridge University Press.

Nation, I.S.P. (2017). The BNC/COCA Level 6 word family lists (Version 1.0.0) [Data file]. Available from http://www.victoria.ac.nz/lals/staff/paul-nation.aspx

Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Boston: Heinle.

Nelson, T. O. (1977). Repetition and depth of processing. *Journal of Verbal Learning and Verbal Behavior, 16*(2), 151–171. https://doi.org/10.1016/S0022-5371(77)80044-3

Newton, J. (2020). Approaches to learning vocabulary inside the classroom. In S. Webb (Ed.), *The Routledge Handbook of Vocabulary* (pp 255-270). New York, New York. doi: 10.4324/9780429291586

Nguyen, C-D., & Boers, F. (2019). The Effect of Content Retelling on Vocabulary Uptake From a TED Talk. *TESOL Quarterly, 53*(1), 5–29. https://doi.org/10.1002/tesq.441

Pavia, N., Webb, S., & Faez, F. (2019). Incidental vocabulary learning through listening to songs. *Studies in Second Language Acquisition*, *41*(4), 745–768. https://doi.org/10.1017/S0272263119000020

Peters, E. (2020) The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. In S. Webb (Ed), Approaches to Learning, Testing and Researching L2 Vocabulary (Vol. 109). John Benjamins Publishing Company.

Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through watching a single episode of L2 television. *Studies in Second Language Acquisition*, *40*(3), 551–557. https://doi.org/10.1017/ S0272263117000407

Puimège, E., Peters, E. (2020). Learning formulaic sequences through viewing L2 television and factors that affect learning. *Studies In Second Language Acquisition*, *42* (3), 525-549. doi: 10.1017/S027226311900055X

Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning, 52*(3), 513–536. https://doi.org/10.1111/1467-9922.00193

Rodgers, M. P. H. (2013). *English language learning through viewing television: An investigation of comprehension, incidental vocabulary acquisition, lexical coverage, attitudes, and captions*. Unpublished PhD thesis, Victoria University, Wellington, New Zealand.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research, 12*, 329–363. https://doi.org/10.1177/1362168808089921

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan.

Swain, M. (1985). Communicative competence: some roles of comprehension input and output in its development. In S. Gass and C. Madden (Eds), *Input in Second Language Acquisition* (pp. 235-253). Rowley, Mass: Newbury House.

Swain, M. (2005). The output hypothesis: theory and research. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (pp. 471-483). Mahwah, N.J.: Lawrence Erlbaum Association

Tegge, F. (2015). *Investigating song-based language teaching and its effect on lexical learning* (Unpublished doctoral dissertation). Victoria University of Wellington, New Zealand. Doi 10.13140/RG.2.1.3159.8560.

Tomczak, E., & Lew, R. (2019). "The Song of words": Teaching multi-word units with songs. *3L: The Southeast Asian Journal of English Language Studies, 25*(4), 16–33.

Treisman. (19645). Monitoring and storage of irrelevant messages in selective attention. *Journal of Verbal Learning and Verbal Behavior*, *3*(6), 449–459. https://doi.org/10.1016/S0022-5371(64)80015-3

Treisman. (1969). Strategies and models of selective attention. *Psychological Review*, *76*(3), 282–299. https://doi.org/10.1037/h0027242

Van Zeeland, H.,&Schmitt, N. (2013). Lexical coverage and L1 and L2 listening comprehension: The same or different from reading comprehension. *Applied Linguistics, 34,* 457–479. doi: 10.1093/applin/ams074

Vidal, K. (2003). Academic listening: A source of vocabulary acquisition? *Applied Linguistics*, *24*(1), 56–89. https://doi.org/10.1093/applin/24.1.56

Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, *61*(1), 219–258. https://doi.org/10.1111/j.1467-9922.2010.00593.x

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, *15*(2), 130–163.

Warner. R. M. (2013). Applied statistics: from bivariate through multivariate techniques (2nd ed.). *SAGE Publications.*

Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition, 27*(1), 33-52. https://doi.org/10.1017/S0272263105050023

Webb, S. (2008). Receptive and productive vocabulary size. *Studies in Second Language Acquisition*, 30(1), 79-95. https://doi.org/10.1017/S0272263108080042

Webb, S. (2009). The effects of receptive and productive learning of word pairs on vocabulary knowledge. *RELC Journal, 40*(3), 360-376. https://doi.org/10.1177/0033688209343854

Webb, S., & Chang, A. C. S. (2015). How does prior word knowledge affect vocabulary learning progress in an extensive reading program*? Studies in Second Language Acquisition, 37,* 651–675

Webb, S., & Nation, P. (2017). *How vocabulary is learned.* Oxford, UK: Oxford University Press.

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. ITL. *International Journal of Applied Linguistics, 168*(1), 34-70.

Werner, V. (2020). "Song-advantage" or "Cost of Singing"? A Research Synthesis of Classroom-based Intervention Studies Applying Lyrics-based Language Teaching (1972–2019). *Journal of Language Teaching and Research, 8* (1), 138-170.

Yanagisawa, A., & Webb, S. (2021). To What Extent Does the Involvement Load Hypothesis Predict Incidental L2 Vocabulary Learning? A Meta-Analysis. *Language Learning, 71*(2), 487–536. https://doi.org/10.1111/lang.12444

Yanagisawa, A., & Webb, S. (2022). Involvement Load Hypothesis Plus: Creating an Improved Predictive Model of Incidental Vocabulary Learning. *Studies in Second Language Acquisition, 44*(5), 1279–1308. https://doi.org/10.1017/S0272263121000577

Yu, X., Boers, F., & Tremblay, P. (2022). Learning multiword items through dictation and dictogloss: How task performance predicts learning outcomes. *Language Teaching Research*, 136216882211172–. https://doi.org/10.1177/13621688221117242

Zhang, X. (2017). Reading–writing integrated tasks, comprehensive corrective feedback, and EFL writing development. *Language Teaching Research*, 21(2), 217–240. https://doi.org/10.1177/1362168815623291

Zou, D., Wang, F. L., Kwan, R., & Xie, H. (2018). Investigating the Effectiveness of Vocabulary Learning Tasks from the Perspective of the Technique Feature Analysis: The Effects of Pictorial Annotations. *In Technology in Education. Innovative Solutions and Practices* (pp. 3–15). Springer Singapore. https://doi.org/10.1007/978-981-13-0008-0_1

Chapter 5: Conclusion

The studies carried out in this research project explored using songs for second/foreign language vocabulary development. The first study replicated the most effective learning condition described by Pavia et al. (2019). In addition, a spaced practice condition was incorporated into the procedure to discover a more efficient and ecologically valid approach for utilizing songs compared to past research. Another difference between Study One of this volume and the previous research was the proficiency level of the participants. The current participant knew approximately 1,660 of the most frequent word families, which was more than three times the participant's knowledge level in the previous research, with approximate knowledge of 430 of the most frequent word families (Pavia et al., 2019. p.750).

The effects of proficiency level on learning outcomes were not used as one of the main dependent measures in this study. Nevertheless, the considerable difference between the current participants' proficiency level and those in the previous research was cautiously taken into account for the interpretation of the current study's outcomes. The results of this study showed that repeated listening (i.e., Five times) to a song could foster the incidental learning of spoken form recognition and form-meaning connection of single words in addition to spoken form recognition of collocations. The learning gains were higher than those found in the previous study (Pavia et al., 2019). Thus, it is reasonable to assume that the variance in results between this study and the 2019 study for mass listening could be attributed to the higher overall proficiency of the participants in this study, as this was the only notable difference between the two studies.

In addition, the comparison of learning outcomes between the spaced and mass listening groups showed that mass practice could be more effective for immediate learning outcomes.

However, based on the results of the delayed posttest, the mass listening condition showed a higher percentage of regression than the spaced listening condition. The analysis of the delayed posttest results showed that there was only one noticeable difference between the mass and spaced listening groups. This difference was observed in the recognition of spoken single words, where the mass listening group outperformed the spaced listening group by a margin of 1.9 words.

The findings in this first study have two main implications. First, they provided further evidence that the incidental vocabulary learning practices proposed by Pavia et al. (2019) can, in fact, foster incidental vocabulary acquisition and that learning gains may be affected by the overall proficiency of the learners. This implies that advanced learners may have an easier time acquiring new vocabulary knowledge. The second implication from this study is that listening to music in a spaced-out manner, similar to how we naturally listen to music, may be just as effective, if not more effective, than listening to music in mass quantities. This is because it leads to longer-lasting learning over time.

The second study of this volume explored the effects of different modes of input when using songs for incidental vocabulary learning. The modes of input included listening only (L), listening while reading the lyrics (LL), listening while singing along, and listening and singing while reading the lyrics (LSL). This study solely focused on the learning of formulaic sequences (FS) as the means to measure vocabulary learning. In addition, considering the presence of receptive and productive learning conditions, word knowledge's receptive and productive aspects were measured separately. The results showed that all learning conditions promoted incidental learning of receptive and productive aspects of FS. However, the LSL group showed significantly higher learning gains than the other learning conditions.

The main implication from the second study in this volume is that using a multimodal approach to language learning from songs can be effective in improving vocabulary acquisition. This approach involves listening, singing, and utilizing lyrics to enhance language learning activities by maximizing learning opportunities through input and output.

The final study in this volume built upon the findings of the first two studies and investigated how retrieval activities can impact the learning of FS when exposed to a specific sequence of input, output, and input again through a song. The results demonstrated that sequential use of input and output in vocabulary learning from songs could foster the development of receptive and productive aspects of FS. Moreover, the added retrieval condition, where the participants are required to produce the target language from memory, had the largest impact on learning gains. Implying that added involvement with song-based activities through recall practices may increase the learner’s attention to the target language resulting in more favourable learning outcomes.

Overall, the findings from the three studies described in this dissertation provided evidence that songs can be considered an invaluable source of language input. However, they were not without limitations and further research is needed. For example, the role of teachers’ perceptions of using songs in classroom instructions was not considered. Studies exploring teachers' self-efficacy beliefs have shown that more efficacious teachers can positively impact their students' learning outcomes (Klassen & Tze, 2014). Thus, it can be speculated that teachers' personal beliefs about their ability to use songs for language education could potentially influence their instructions' effectiveness when using songs.

Tegge (2018) examined teachers' beliefs about using songs in L2 teaching through a series of surveys. The findings revealed that teachers found songs to be a useful tool for

second/foreign language teaching. Most of the respondents reported that they used songs to create a motivating learning environment for students. They recognized songs as an authentic language and cultural source, allowing their students to practice listening comprehension. Nevertheless, many teachers refrained from using songs and identified several challenges preventing them from using songs. Among the challenges, they reported that they had difficulty finding suitable songs, and song use required lots of preparation time, which outweighed the learning outcomes. Sevik (2011) found similar reports from 52 Turkish state primary school EFL teachers. Teachers in this context found songs to be pedagogically valuable for teaching English as a foreign language. 63.5% of the respondents reported using songs as much as possible. However, 44.3% of the teachers reported that a lack of resources and difficulty finding songs for different aspects of language teaching discouraged them from using songs regularly. These reports were collected from experienced teachers who had taken it upon themselves to use songs as a source of language input without formal training on how and when to use songs for language education. Based on these findings, researchers speculated that if teachers received formal training on using songs and had access to pre-existing teaching materials that clearly outline the learning goals from different songs, they would become more confident in using songs effectively. Subsequently, they would use songs for language education more frequently (Tegge, 2018; Sevik, 2011). While these speculations are logical, further empirical investigation is needed to explain the relationship between teachers' beliefs, song use, and students' language learning gains.

5.1 References

Klassen, R.M., & Tze, V.M.C. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. Educational Research Review, 12, 59 – 76. https://doi.org/10.1016/j.edurev.2014.06.001

Pavia, N., Webb, S., & Faez, F. (2019). Incidental vocabulary learning through listening to songs. *Studies in Second Language Acquisition*, *41*(4), 745–768. https://doi.org/10.1017/S0272263119000020

Sevik, M. (2011). Teacher views about using songs in teaching English to young learners. *Educational Research and Review, 6*(21), 1027-1035. DOI: 10.5897/ERR11.250

Tegge, F. (2018). Pop songs in the classroom: Time-filler or teaching tool? *ELT Journal, 72*, 274–284. doi: https://doi.org/10.1093/elt/ccx071

Appendices

Appendix A: Western University Ethics Approval

Western Research

Date: 4 October 2021

To: Dr. Farahnaz Faez

Project ID: 118862

Study Title: Language Learning Through Songs

Short Title: Language and Songs

Application Type: NMREB Initial Application

Review Type: Delegated

Full Board Reporting Date: 05/Nov/2021

Date Approval Issued: 04/Oct/2021 16:19

REB Approval Expiry Date: 04/Oct/2022

Dear Dr. Farahnaz Faez

The Western University Non-Medical Research Ethics Board (NMREB) has reviewed and approved the WREM application form for the above mentioned study, as of the date noted above. NMREB approval for this study remains valid until the expiry date noted above, conditional to timely submission and acceptance of NMREB Continuing Ethics Review.

This research study is to be conducted by the investigator noted above. **All other required institutional approvals and mandated training must also be obtained prior to the conduct of the study.**

Documents Approved:

Document Name	Document Type	Document Date	Document Version
Vocabulary Level Test (English-Thai)	Paper Survey	23/Mar/2021	
The updated Vocabulary Levels Test	Paper Survey	23/Mar/2021	
Study 1- Vocabulary Knowledge Test (Pavia et al.2019)	Paper Survey	23/Mar/2021	
Study 2- Vocabulary KnowledgeTest	Paper Survey	23/Mar/2021	
Interview Questions	Interview Guide	23/Mar/2021	
Participant Number	Paper Survey	05/Aug/2021	1
Script for recruiting Student-Participants (Control groups)	Oral Script	05/Aug/2021	1
Script for recruiting Student-Participants (Experimental Groups)	Oral Script	05/Aug/2021	
Script for recruiting Teacher-Participants	Oral Script	05/Aug/2021	1
Demographic information sheet	Paper Survey	05/Aug/2021	1
Script for obtaining verbal consent	Verbal Consent/Assent	20/Mar/2021	1
Letter of permission School director Thailand	Written Consent/Assent	22/Mar/2021	1
Study 1 & 2- Letter of Information: Consent form - Control Groups	Written Consent/Assent	16/Sep/2021	1
Study 1- Letter of Information - E1	Written Consent/Assent	16/Sep/2021	1
Study 1- Letter of Information - E2	Written Consent/Assent	16/Sep/2021	1
Study 2- Letter of Information - E1	Written Consent/Assent	16/Sep/2021	1

Study 2- Letter of Information - E2	Written Consent/Assent	16/Sep/2021	1
Study 2- Letter of Information - E3	Written Consent/Assent	16/Sep/2021	1
Study 2- Letter of Information - E4	Written Consent/Assent	16/Sep/2021	1
Study 2- Letter of Information - E5 & E6	Written Consent/Assent	16/Sep/2021	1
Letter of Information Teacher-Participants	Written Consent/Assent	27/Sep/2021	1

Documents Acknowledged:

Document Name	Document Type	Document Date	Document Version
Study 1-Timeline and Procedure	Supplementary Tables/Figures	23/Mar/2021	
Study 2-Timeline & Procedure	Supplementary Tables/Figures	23/Mar/2021	
Study 1 - Target Vocabulary	Supplementary Tables/Figures	23/Mar/2021	
Study 2- Target Vocabulary	Supplementary Tables/Figures	23/Mar/2021	
Study 1- Song Lyrics	Supplementary Tables/Figures	23/Mar/2021	
Study 2 - Song Lyrics	Supplementary Tables/Figures	23/Mar/2021	

No deviations from, or changes to the protocol should be initiated without prior written approval from the NMREB, except when necessary to eliminate immediate hazard(s) to study participants or when the change(s) involves only administrative or logistical aspects of the trial.

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario. Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB. The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000941.

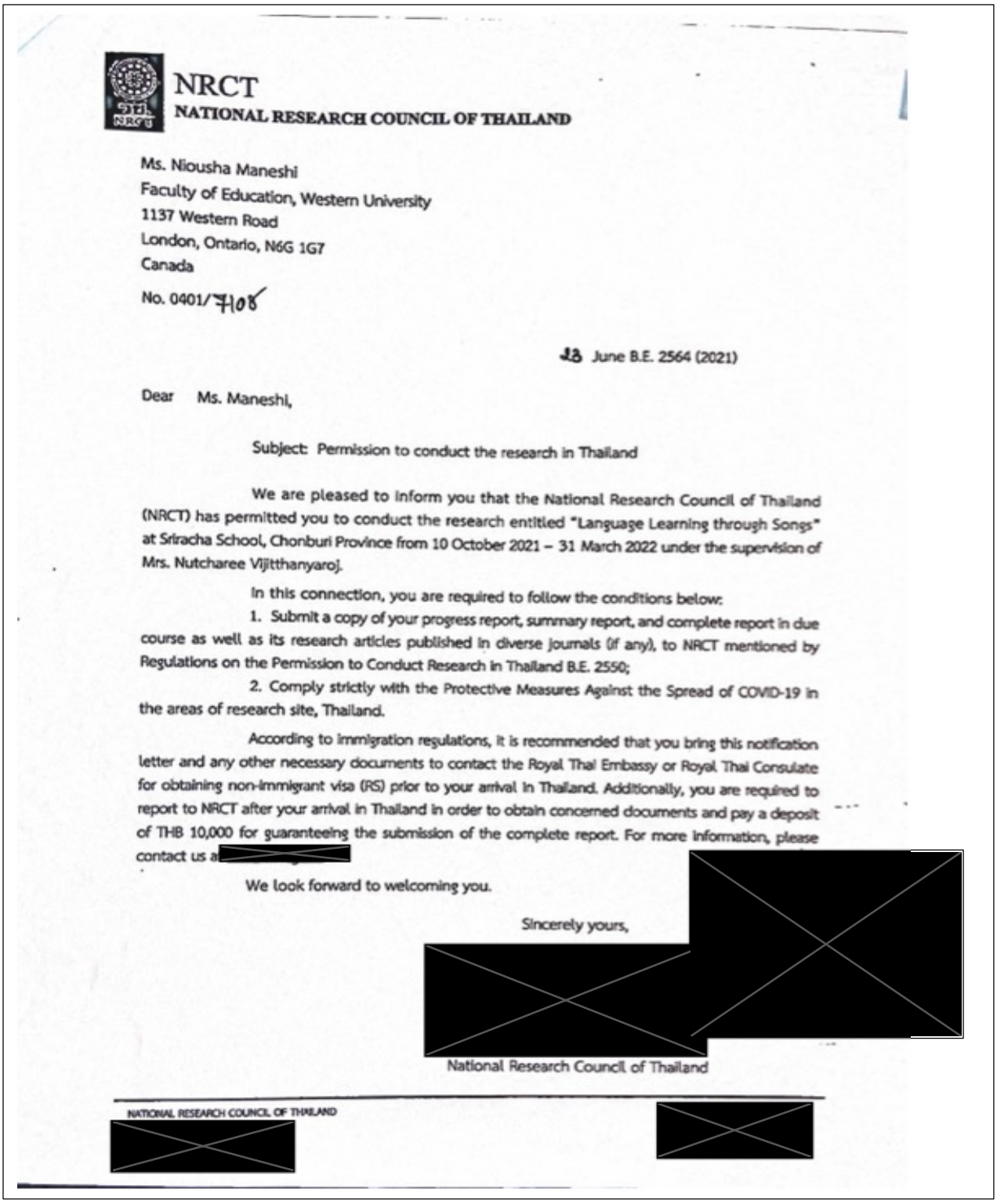
Please do not hesitate to contact us if you have any questions.

Sincerely,

Ms. Katelyn Harris , Research Ethics Officer on behalf of Dr. Randal Graham, NMREB Chair

Note: This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).

Appendix B: NRCT Approval



NRCT
NATIONAL RESEARCH COUNCIL OF THAILAND

Ms. Niousha Maneshi
Faculty of Education, Western University
1137 Western Road
London, Ontario, N6G 1G7
Canada

No. 0401/7108

18 June B.E. 2564 (2021)

Dear Ms. Maneshi,

Subject: Permission to conduct the research in Thailand

We are pleased to inform you that the National Research Council of Thailand (NRCT) has permitted you to conduct the research entitled "Language Learning through Songs" at Sriracha School, Chonburi Province from 10 October 2021 – 31 March 2022 under the supervision of Mrs. Nutcharee Vijitthanyaroj.

In this connection, you are required to follow the conditions below:

1. Submit a copy of your progress report, summary report, and complete report in due course as well as its research articles published in diverse journals (if any), to NRCT mentioned by Regulations on the Permission to Conduct Research in Thailand B.E. 2550;
2. Comply strictly with the Protective Measures Against the Spread of COVID-19 in the areas of research site, Thailand.

According to immigration regulations, it is recommended that you bring this notification letter and any other necessary documents to contact the Royal Thai Embassy or Royal Thai Consulate for obtaining non-immigrant visa (RS) prior to your arrival in Thailand. Additionally, you are required to report to NRCT after your arrival in Thailand in order to obtain concerned documents and pay a deposit of THB 10,000 for guaranteeing the submission of the complete report. For more information, please contact us a[redacted]

We look forward to welcoming you.

Sincerely yours,

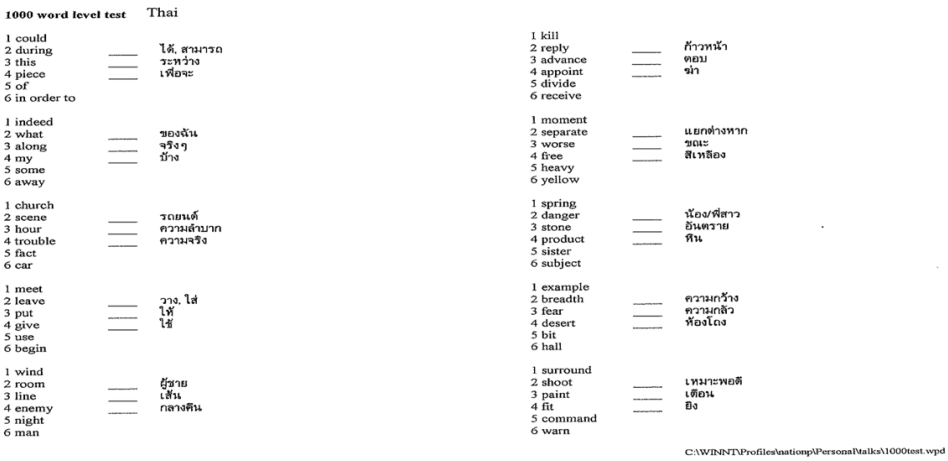
[redacted]

National Research Council of Thailand

NATIONAL RESEARCH COUNCIL OF THAILAND
[redacted]

Appendix C: Bilingual (Thai-English) Vocabulary Level Test

1000 Word level test: English and Thai



1000 word level test Thai

1 could
2 during ____ ได้, สามารถ
3 this ____ ระหว่าง
4 piece ____ เพื่อจะ
5 of
6 in order to

1 indeed
2 what ____ ของฉัน
3 along ____ จริงๆ
4 my ____ บ้าง
5 some
6 away

1 church
2 scene ____ รถยนต์
3 hour ____ ความลำบาก
4 trouble ____ ความจริง
5 fact
6 car

1 meet
2 leave ____ วาง, ใส่
3 put ____ ให้
4 give ____ ใช้
5 use
6 begin

1 wind
2 room ____ ผู้ชาย
3 line ____ เส้น
4 enemy ____ กลางคืน
5 night
6 man

1 kill
2 reply ____ ก้าวหน้า
3 advance ____ ตอบ
4 appoint ____ ฆ่า
5 divide
6 receive

1 moment
2 separate ____ แยกต่างหาก
3 worse ____ ขณะ
4 free ____ สีเหลือง
5 heavy
6 yellow

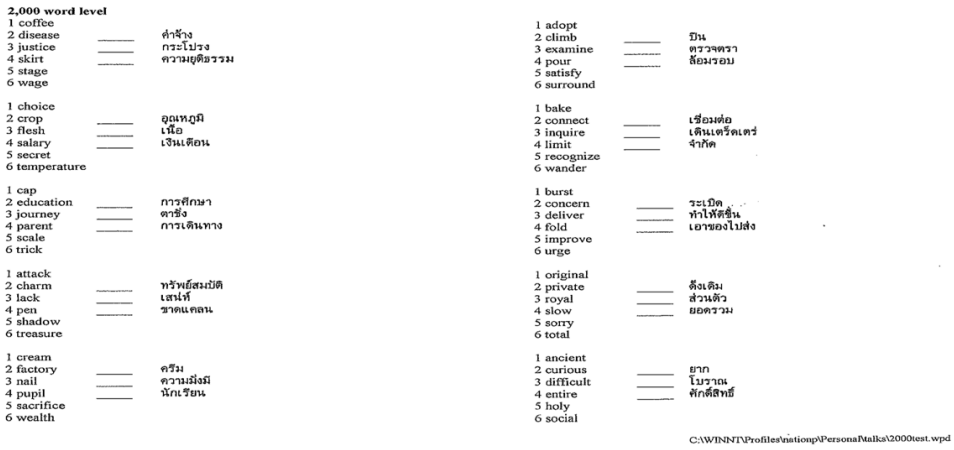
1 spring
2 danger ____ น้อง/พี่สาว
3 stone ____ อันตราย
4 product ____ หิน
5 sister
6 subject

1 example
2 breadth ____ ความกว้าง
3 fear ____ ความกลัว
4 desert ____ ห้องโถง
5 bit
6 hall

1 surround
2 shoot ____ เหมาะพอดี
3 paint ____ เตือน
4 fit ____ ยิง
5 command
6 warn

C:\WINNT\Profiles\nationp\Personal\talks\1000test.wpd

2000 Word level test: English and Thai



2,000 word level
1 coffee
2 disease ____ ค่าจ้าง
3 justice ____ กระโปรง
4 skirt ____ ความยุติธรรม
5 stage
6 wage

1 choice
2 crop ____ อุณหภูมิ
3 flesh ____ เนื้อ
4 salary ____ เงินเดือน
5 secret
6 temperature

1 cap
2 education ____ การศึกษา
3 journey ____ ตาชั่ง
4 parent ____ การเดินทาง
5 scale
6 trick

1 attack
2 charm ____ ทรัพย์สมบัติ
3 lack ____ เสน่ห์
4 pen ____ ขาดแคลน
5 shadow
6 treasure

1 cream
2 factory ____ ครีม
3 nail ____ ความมั่งมี
4 pupil ____ นักเรียน
5 sacrifice
6 wealth

1 adopt
2 climb ____ ปีน
3 examine ____ ตรวจตรา
4 pour ____ ล้อมรอบ
5 satisfy
6 surround

1 bake
2 connect ____ เชื่อมต่อ
3 inquire ____ เดินเตร็ดเตร่
4 limit ____ จำกัด
5 recognize
6 wander

1 burst
2 concern ____ ระเบิด
3 deliver ____ ทำให้ตื่นเต้น
4 fold ____ เอาของไปส่ง
5 improve
6 urge

1 original
2 private ____ ดั้งเดิม
3 royal ____ ส่วนตัว
4 slow ____ ยอดรวม
5 sorry
6 total

1 ancient
2 curious ____ ยาก
3 difficult ____ โบราณ
4 entire ____ พักผ่อนสิทธิ์
5 holy
6 social

C:\WINNT\Profiles\nationp\Personal\talks\2000test.wpd

Appendix D: The updated Vocabulary Levels Test (Webb, Sasao, & Ballance, 2017)

This is test that looks at how well you know useful English words. Put a check under the word that goes with each meaning. Here is an example.

	game	island	mouth	movie	song	yard
land with water all around it						
part of your body used for eating and talking						
piece of music						

It should be answered in the following way.

	game	island	mouth	movie	song	yard
land with water all around it		✓				
part of your body used for eating and talking			✓			
piece of music					✓	

1,000 Word Level

	choice	computer	garden	photograph	price	week
cost						
picture						
place where things grow outside						

	eye	father	night	van	voice	year
body part that sees						
parent who is a man						
part of the day with no sun						

	center	note	state	tomorrow	uncle	winter
brother of your mother or father						
middle						
short piece of writing						

	box	brother	horse	hour	house	plan
family member						
sixty minutes						
way of doing things						

	animal	bath	crime	grass	law	shoulder

green leaves that cover the ground						
place to wash						
top end of your arm						

	drink	educate	forget	laugh	prepare	suit
get ready						
make a happy sound						
not remember						

	check	fight	return	tell	work	write
do things to get money						
go back again						
make sure						

	bring	can	reply	stare	understand	wish
say or write an answer to somebody						
carry to another place						
look at for a long time						

	alone	bad	cold	green	loud	main
most important						
not good						
not hot						

	awful	definite	exciting	general	mad	sweet
certain						
usual						
very bad						

2,000 Word Level

	coach	customer	feature	pie	vehicle	weed
important part of something						
person who trains members of sports teams						
unwanted plant						

	average	discipline	knowledge	pocket	trap	vegetable
food grown in gardens						

information which a person has						
middle number						

	circle	justice	knife	onion	partner	pension
round shape						
something used to cut food						
using laws fairly						

	cable	section	sheet	site	staff	tank
part						
place						
something to cover a bed						

	apartment	cap	envelope	lawyer	speed	union
cover for letters						
kind of hat						
place to live inside a tall building						

	argue	contribute	quit	seek	vote	wrap
cover tightly and completely						
give to						
look for						

	avoid	contain	murder	search	switch	trade
have something inside						
look for						
try not to do						

	bump	complicate	include	organize	receive	warn
get something						
hit gently						
have as part of something						

	available	constant	electrical	medical	proud	super
feeling good about what you have done						
great						
happening all the time						

	environmental	junior	pure	rotten	smooth	wise
bad						
not rough						
younger in position						

3,000 Word Level

	angle	apology	behavior	bible	celebration	portion
actions						
happy occasion						
statement saying you are sorry						

	anxiety	athlete	counsel	foundation	phrase	wealth
combination of words						
guidance						
large amount of money						

	agriculture	conference	frequency	liquid	regime	volunteer
farming						
government						
person who helps without payment						

	asset	heritage	novel	poverty	prosecution	suburb
having little money						
history						
useful thing						

	audience	crystal	intelligence	outcome	pit	welfare
ability to learn						
deep place						
people who watch and listen						

	consent	enforce	exhibit	retain	specify	target
agree						
say clearly						
show in public						

	accomplish	capture	debate	impose	proceed	prohibit
catch						

go on						
talk about what is correct						

	absorb	decline	exceed	link	nod	persist
continue to happen						
goes beyond the limit						
take in						

	approximate	frequent	graphic	pale	prior	vital
almost exact						
earlier						
happening often						

	consistent	enthusiastic	former	logical	marginal	mutual
not changing						
occurring earlier in time						
shared						

Appendix E: Lyrics of "Die a Happy Man" Thomas Rhett (2015)

Retrieved from https://www.azlyrics.com/lyrics/thomasrhett/dieahappyman.html

Baby, last night was hands down
One of the best nights
That I've had no doubt
Between the bottle of wine
And the look in your eyes and the Marvin Gaye
Then we danced in the dark under September stars in the pourin' rain

And I know that I can't ever tell you enough
That all I need in this life is your crazy love

If I never get to see the Northern lights
Or if I never get to see the Eiffel Tower at night
Oh, if all I got is your hand in my hand
Baby, I could die a happy man

Happy man, baby
Mmm

Baby, that red dress brings me to my knees
Oh, but that black dress makes it hard to breathe
You're a saint, you're a Goddess,
The cutest, the hottest,
A masterpiece
It's too good to be true,
Nothing better than you
In my wildest dreams

And I know that I can't ever tell you enough
That all I need in this life is your crazy love

If I never get to see the Northern lights
Or if I never get to see the Eiffel Tower at night
Oh, if all I got is your hand in my hand
Baby, I could die a happy man, yeah

I don't need no vacation,
No fancy destination
Baby, you're my great escape
We could stay at home,
Listen to the radio
Or dance around the fireplace

And if I never get to build my mansion in Georgia
Or drive a sports car up the coast of California
Oh, if all I got is your hand in my hand
Baby, I could die a happy man

Baby, I could die a happy man
Oh, I could die a happy man
You know I could girl
I could die, I could die a happy man

Appendix F : Vocabulary Knowledge Test (Pavia et al., 2019)

Text of the audio recorded test. All questions measure receptive knowledge.

A. Form Recognition Test:

The participant sees on paper:

1. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say:

'Number one [1sec.]

a) Die [2sec.]

b) Sef [2sec.]

c) Huj [4sec.]'

2. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say:

'Number two [1sec.]

a) Conbex [2sec.]

b) Enough [2sec.]

c) Shaste [2 sec.]

3. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say:

'Number three' [1sec.]

a) Copaque [2sec.]

b) Between [2sec.]

c) Sluster [2 sec.]

4. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say:

‘Number four’ [1sec.]

a) Build [2sec.]

b) Klade [2sec.]

c) Tover [2 sec.]

5. A B C I don’t remember any of these (In Thai)

At the same time the participant hears the teacher say:

‘Number five’ [1sec.]

a) Last [2sec.]

b) Gavo [2sec.]

c) Hane [4sec.]’

6. A B C I don’t remember any of these (In Thai)

At the same time the participant hears the teacher say:

‘Number six’ [1sec.]

a) Meaj [2sec.]

b) Flob [2sec.]

c) True [2 sec.]

7. A B C I don’t remember any of these (In Thai)

At the same time the participant hears the teacher say:

‘Number seven’ [1sec.]

a) Kopal [2sec.]

b) Botan [2sec.]

c) Under [2 sec.]

8. A B C I don’t remember any of these (In Thai)

At the same time the participant hears the teacher say:

‘Number eight’ [1sec.]

a) Fancy [2sec.]

b) Moyal [2sec.]

d) Dlurt [4sec.]'

9. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say:

'Number nine' [1sec.]

a) Terish [2sec.]

b) Escape [2sec.]

c) Kigest [2 sec.]

10. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say:

'Number ten' [1sec.]

a) Banzu [2sec.]

b) Rupon [2sec.]

c) Coast [2 sec.]

11. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say:

'Number eleven' [1sec.]

a) Dark [2sec.]

b) Hiwi [2 sec.]

c) Gent [4sec.]'

12. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say:

'Number tweleve'[1sec.]

a) Naby [2sec.]

b) Star [2sec.]

c) Tane [4sec.]'

13. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say:

'Number thirteen [1sec.]

a) Knee [2sec.]

b) Fres [2sec.]

c) Kade [2 sec.]

14. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say:

'Number fourteen [1sec.]

a) Pacon [2sec.]

b) Gumly [2 sec.]

c) Saint [4sec.]'

15. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say:

'Number fifteen [1sec.]

a) Conjudate [2sec.]

b) Fomposure [2sec.]

c) Fireplace [4sec.]'

16. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say:

'Number sixteen [1sec.]

a) Bustifiable [2sec.]

b) Fonozygotic [2 sec.]

c) Destination [4sec.]'

17. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say:

'Number seventeen [1sec.]

a) Porfame [2sec.]

b) Mansion [2sec.]

c) Shobalt [2 sec.]

18. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say:

'Number eighteen [1sec.]

a) Vacation [2sec.]

b) Prolinic [2 sec.]

c) Infecent [4sec.]'

19. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say:

'Number nineteen [1sec.]

a) Kathematize [2sec.]

b) Masterpiece [2 sec.]

c) Brequenciesf [4sec.]'

B. Form Meaning Connection

Translation

The participants will hear the targeted word and asked to select the correct answer they see on paper which will be in Thai.

1. Students will hear "Last"

a) Single

b) Since

c) Last

d) I don't know

2. Students will hear "Die"

a) True

b) Drive

c) Die

d) I don't know

3. Students will hear "Enough"

a) Nothing

b) Enough

c) Escape

d) I don't know

4. Students will hear "Between"

a) Build

b) Between

c) Stay

d) I don't know

5. Students will hear "True"

a) True

b) Dark

c) Great

d) I don't know

6. Students will hear "Fancy"
    a) Life
    b) Vacation
    c) Fancy
    d) I don't know
7. Students will hear "Dark"
    a) Dark
    b) Tower
    c) Saint
    d) I don't know
8. Students will hear "Under"
    a) Under
    b) Drive
    c) Mansion
    d) I don't know
9. Students will hear "Coast"
    a) Lights
    b) Vacation
    c) Coast
    d) I don't know
10. Students will hear "Knee"
    a) Knee

b) Bottle

c) Enough

d) I don't know

11. Students will hear "Saint"

a) Crazy

b) Saint

c) Escape

d) I don't know

12. Students will hear "Destination"

a) Goddess

b) Fireplace

c) Destination

d) I don't know

13. Students will hear "Masterpiece"

a) Wine

b) Coast

c) Masterpiece

d) I don't know

14. Students will hear "Vacation"

a) Vacation

b) Die

c) Wildest

d) I don't know

15. Students will hear “Escape”
    a) Fancy
    b) Escape
    c) Nothing
    d) I don’t know
16. Students will hear “Mansion”
    a) Enough
    b) Breathe
    c) Mansion
    d) I don’t know
17. Students will hear “Star”
    a) Star
    b) Dark
    c) Eyes
    d) I don’t know
18. Students will hear “Fireplace”
    a) Life
    b) Fireplace
    c) Masterpiece
    d) I don’t know
19. Students will hear “Build”
    a) Die
    b) Build

c) Aches

d) I don't know

C. Collocations

Matching:

Choose the correct match to form word partnership.

Example (*This example will be written on the class board and explained prior to starting this section of the test)*

A) Sing a song B) Sing a food C) Sing a ball D) I don't remember any of these (in Thai)

The participant sees on paper:

1 . A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say: 'Number one [1sec.]

a) Bottle of dress [2sec.]

b) Bottle of wine [2sec.]

c) Bottle of crazy [2sec.]

2. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say: 'Number two [1sec.]

a) Pouring rain [2sec.]

b) Pouring baby [2sec.]

c) Pouring need [2sec.]

3. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say: 'Number three [1sec.]

a) Northern lights [2sec.]

b) Northern true [2sec.]

c) Northern cutest [2sec.]

4. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say: 'Number four [1sec.]

a) Listen to the life [2sec.]

b) Listen to the down [2sec.]

c) Listen to the radio [2sec.]

5. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say: 'Number five [1sec.]

a) Sports eyes [2sec.]

b) Sports car [2sec.]

c) Sports baby [2sec.]

6. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say: 'Number six [1sec.]

a) Northern Doubt [2sec.]

b) Red doubt [2sec.]

c) No doubt [2sec.]

7. A B C I don't remember any of these (In Thai)

At the same time the participant hears the teacher say: 'Number seven [1sec.]

a) Wildest down [2sec.]

b) Wildest dreams [2sec.]

c) Wildest black [2sec.]

Appendix G: Vocabulary Knowledge test: Participant Form (Pavia et al., 2019)

คำถามนี้จะเป็นตัววัดความสามารถในการรับรู้
แบบทดสอบการรับรู้: จงเลือกคำตอบที่ถูกต้องจากเสียงที่ได้ฟัง

1. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
2. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
3. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
4. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
5. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
6. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
7. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
8. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
9. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
10. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
11. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
12. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
13. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
14. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
15. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
16. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐

17. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐

18. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐

19. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐

B. การแปล: จงเลือกคำที่มีความหมายที่ถูกต้องจากเสียงที่ได้ฟัง

1. A) **โสด** B) ตั้งแต่ C) สุดท้าย D) **ไม่ทราบ**

2. A) ความจริง B) ขับขี่ C) เสียชีวิต D) **ไม่ทราบ**

3. A) **ไม่มีอะไร / ไร้ความหมาย** B) เพียงพอ C) **หลบหนี** D) **ไม่ทราบ**

4. A) ก่อสร้าง B) ระหว่าง C) อาศัยอยู่ D) **ไม่ทราบ**

5. A) ความจริง B) มืด C) **ดีเยี่ยม** D) **ไม่ทราบ**

6. A) ชีวิต B) **ช่วงพักผ่อน** C) จินตนาการ / หรูหรา D) **ไม่ทราบ**

7. A) มืด B) **หอคอย** C) **นักบุญ** D) **ไม่ทราบ**

8. A) **ภายใต้ / ข้างล่าง** B) ขับขี่ C) คฤหาสน์ D) **ไม่ทราบ**

9. A) ส่องแสง B) ช่วงพักผ่อน C) ชายฝั่ง D) **ไม่ทราบ**

10. A) **หัวเข่า** B) ขวด C) เพียงพอ D) **ไม่ทราบ**

11. A) บ้าคลั่ง B) นักบุญ C) **หลบหนี** D) **ไม่ทราบ**

12. A) เทพธิดา B) เตาผิง C) **จุดหมายปลายทาง** D) **ไม่ทราบ**

13. A) **เหล้าองุ่น** B) ชายฝั่ง C) งานชิ้นเอก D) **ไม่ทราบ**

14. A) ช่วงพักผ่อน B) **เสียชีวิต** C) ดุร้าย/ป่าเถื่อน D) **ไม่ทราบ**

15. A) จินตนาการ/หรูหรา	B) **หลบหนี**	**C) ไม่มีอะไร/ไร้ความหมาย**	D) **ไม่ทราบ**
16. A) เพียงพอ	B) **ลมหายใจ**	C) **คฤหาสน์**	D) **ไม่ทราบ**
17. A) **ดวงดาว**	B) **ความมืด**	C) ดวงตา	D) **ไม่ทราบ**
18. A) ชีวิต	B) **เตาผิง**	C) **งานชิ้นเอก**	D) **ไม่ทราบ**
19. A) เสียชีวิต	B) ก่อสร้าง	C) อาการปวด	D) **ไม่ทราบ**

C. จับคู่**: จงจับคู่ของคำที่ถูกต้องจากเสียงที่ได้ฟัง**

ตัวอย่าง**:** (A) ■ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐

1. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
2. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
3. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
4. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
5. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
6. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐
7. (A) ☐ (B) ☐ (C) ☐ (D) **ฉันจำอะไรไม่ได้เลย** ☐

Appendix H:Lyrics of "Circles" Post Malone (2019)

Retrieved from https://www.youtube.com/PostMalone/Circles

Oh, oh, oh-oh

Oh, oh, oh-oh

Oh, oh, oh-oh, oh-oh, oh-oh

We couldn't turn around till we were upside down

I'll be the bad guy now, but know I ain't too proud

I couldn't be there even when I try

You don't believe it, we do this every time

Seasons change and our love went cold

Feed the flame 'cause we can't let go

Run away, but we're running in circles

Run away, run away

I dare you to do something

I'm waiting on you again, so I don't take the blame

Run away, but we're running in circles

Run away, run away, run away

Let go, I got a feeling that it's time to let go

I said so, I knew that this was doomed from the get-go

You thought that it was special, special

But it was just a special. The special

And I still hear the echoes (The echoes)

I got a feeling that it's time to let it go, let it go

Seasons change and our love went cold

Feed the flame 'cause we can't let go

Run away, but we're running in circles

Run away, run away

I dare you to do something

I'm waiting on you again, so I don't take the blame

Run away, but we're running in circles

Run away, run away, run away

Maybe you don't understand what I'm going through

It's only me, what you got to lose?

Make up your mind, tell me, what are you gonna do?

It's only me, let it go

Seasons change and our love went cold

Feed the flame 'cause we can't let go

Run away, but we're running in circles

Run away, run away

I dare you to do something

I'm waiting on you again, so I don't take the blame

Run away, but we're running in circles

Run away, run away, run away

Appendix I: Vocabulary Knowledge Test (Study 2 & 3)

Productive Vocabulary knowledge test.

Fill in the blank to complete the poem. (To be Translated)

We couldn't …**1**…..around till we were …..**2**….. down	1……
I'll be the bad guy now, but know I ain't too proud	2……
I couldn't be there even when I tried	3……
You don't believe it, we do this **…..3**…. time	4……
	5……
Seasons …..**4**….. and our love went cold	6……
…**5**…. the flame 'cause we can't …**6**.. go	7……
Run …..**7**…., but we're running in ….**8**….	8……
I'm waiting ….**9**… you again, so I don't ….**10**… the blame	9……
	10……
I ….**11**… a feeling that it's time	11……
I said so, I knew it was doomed from the …**12**…	12……
	13……
Maybe you don't ….**13**…what I'm going…**14**…	14……
It's only me, what you got to…**15**...?	15……
Make up your…**16**…, tell me, what are you gonna do?	16……

Receptive Vocabulary Knowledge Test

Choose the correct word to complete the sentences.

1. ……around.	A)Turn	B) Know	C) Change	D) ไม่ทราบ
2……..down.	A) To	B) Upside	C) Through	D) ไม่ทราบ
3.Waiting -------you again.	A) On	B) Up	C) In	D) ไม่ทราบ
4…….Time.	A) Every	B) Even	C) Maybe	D) ไม่ทราบ
5.Seasons--------.	A) Dare	B) Run	C) Change	D) ไม่ทราบ
6. -------the flame.	A) Change	B) Feed	C) Hear	D) ไม่ทราบ
7.-------go	A) Know	B) Let	C) Got	D) ไม่ทราบ
8. Run--------.	A) Every	B) Only	C) Away	D) ไม่ทราบ
9. we’re running in ---------.	A) Mind	B) Something	C) Circles	D) ไม่ทราบ
10. I ….the blame.	A) Go	B) Take	C) Run	D) ไม่ทราบ
11. I …..a feeling.	A) Change	B) Doomed	C) Got	D) ไม่ทราบ
12. I….so.	A) Lose	B) Said	C) Went	D) ไม่ทราบ
13. From the ------.	A) Circle	B) Get-go	C) Love	D) ไม่ทราบ
14. I don’t------.	A) Got	B) Dare	C) Understand	D) ไม่ทราบ
15. What I’m going-----.	A) On	B) Upside	C) Through	D) ไม่ทราบ
16. Make up your -----.	A) Mind	B) Cold	C) Circles	D) ไม่ทราบ

Curriculum Vitae

Name: Niousha Pavia

Post-secondary Education and Degrees:

The University of Western Ontario
London, Ontario, Canada
2009-2014 B.A.

The University of Western Ontario
London, Ontario, Canada
2015-2017 M.A.

The University of Western Ontario
London, Ontario, Canada
2017- 2023 Ph.D. (To be completed)

Honours and Awards:

The University of Western Ontario
Graduate Excellence in Teaching Award
2019, 2023

Related Work Experience

Graduate Course Instructor
The University of Western Ontario
Teaching listening and speaking skills (Course#GRADEDU9312)
Jan1, 2023- April 30, 2023
Teaching the Four Skills (Course# GRADEDU9304)
Sep 1, 2019 – Dec 31, 2019
Jan 1, 2018 – Apr 30, 2018

Editorial Assistant
TESL Canada Journal
May 1, 2020 – May 1, 2022

TESOL Instructor & Program Coordinator
Explore and Teach Asia (Xplore Asia), Thailand
Sep-Oct 2018
Sep-Oct 2016

Highschool English Teacher
Kuengnai-Pittayakarn School, Thailand
2014-2015

Publications & Conference Presentation:

- Faez, F., Martini, Juliane., & Pavia, N. (2022). Research methods in applied linguistics education: current considerations, recent innovations, and future directions. *ELT Journal.* doi: https://doi.org/10.1093/elt/ccab091
- Pavia, N., Webb, S., & Faez, F. (2019). Incidental vocabulary learning through listening to songs. *Studies in Second Language Acquisition*, 1-24. doi:10.1017/S0272263119000020.
- Maneshi, N. (2017). Incidental vocabulary learning through listening to songs. *Electronic Thesis and Dissertation Repository*. 4783. http://ir.lib.uwo.ca/etd/4783
- American Association for Applied Linguistics (AAAL): Individual paper presentation at the 2018 conference.
- Western University, Faculty of Education: Invitation to speak in the 2017, Symposium on Teaching and Learning Vocabulary in Another Language.