MA Thesis Proposal · Tunghai University, November 2026

Song-Integrated Instruction: Productive Vocabulary Learning and Retention among Young EFL Learners in Taiwan. Approved at the proposal hearing in November 2026. Supervisor: Prof. Chen Meihua. This excerpt shows the title page, the abstract, the opening of the introduction and the research questions.

歌曲融入教學對國小學童英語字彙學習與記憶之影響

Song-Integrated Instruction: Productive Vocabulary Learning and Retention among Young EFL Learners in Taiwan

by

法拉 Farah M.

PROPOSAL

Presented to the Committee

TUNGHAI UNIVERSITY

November 2026

中華民國一百一十五年十一月

© 2026 Farah M. · Portfolio sample for review only · Not for reuse or redistribution · farah.m.academic@gmail.com

Abstract

Abstract

Studies have reported benefits of songs for English vocabulary learning, but the strength of that evidence depends on how learning is measured and what songs are compared with. Several studies have relied on recognition tasks or on tests given soon after teaching, songs have shown no advantage over choral repetition, and a meta-analysis has found smaller benefits for children than for adults. This study asks a harder question: when children must produce the words themselves, weeks after instruction, does a song add anything to the flashcard teaching their class already uses?

The study uses a within-participants design with the 15 learners of one Grade 4 class at a private English language center in Taiwan. Twenty target words from one lesson of the class curriculum were pre-tested, matched in pairs on part of speech, frequency, number of syllables, concreteness, and pre-test score, and assigned to the two conditions at random within each pair. Each learner is therefore taught 10 words through song-integrated instruction and 10 through flashcard instruction, in the same six sessions and for the same amount of time. The researcher teaches both conditions. The song's lyrics and audio were generated with AI and revised by the researcher. A pilot study with the same class shaped the test tasks and the procedure for selecting words.

Productive knowledge is measured four times: before instruction, immediately after it, and 2 and 7 weeks after the immediate post-test. In form recall, learners read a short English definition and write the word. In sentence production, they write a sentence that shows the word's meaning. Two raters score every paper independently, and their scores are averaged. One week after the immediate post-test, a short questionnaire in Chinese asks learners how they experienced the two methods.

Descriptive and inferential analyses are used. For each section of the test, Wilcoxon signed-rank tests compare each learner's scores on the two word sets, with

© 2026 Farah M. · Portfolio sample for review only · Not for reuse or redistribution · farah.m.academic@gmail.com

Abstract (continued)

the rank-biserial correlation as the effect size. For retention, the tests also compare how much of each word set learners lost between the immediate and delayed post-tests. Questionnaire ratings are summarized, and the open answers are analyzed thematically. Parental consent was obtained through the language center, participation is voluntary, and the results do not affect learners' grades.

The study makes two contributions. First, it tests whether a song adds to an established vocabulary technique when the outcome is productive knowledge weeks after teaching, not recognition soon after it. Second, it offers a classroom procedure, refined through a pilot study, for comparing two teaching methods within one class of young learners.

Keywords: song-integrated instruction, productive vocabulary knowledge, vocabulary retention, flashcards, young EFL learners, Taiwan

© 2026 Farah M. · Portfolio sample for review only · Not for reuse or redistribution · farah.m.academic@gmail.com

Chapter One · Background of the Study

CHAPTER ONE

INTRODUCTION

1.1 Background of the Study

In the private English language center where this study takes place, new words are taught through teacher explanation, flashcard review, and workbook exercises. Flashcard review is an established technique for deliberate vocabulary learning (Nation, 2013; Webb & Nation, 2017), and in the participating class it produces good results on unit tests. Unit tests, however, show what learners can do while the words are still fresh. They say little about whether learners can still recall and use the words once the unit is over, and recalling and using a word is harder than recognizing it: receptive knowledge generally develops ahead of productive knowledge (Nation, 2013).

Songs are the other familiar vocabulary tool in young learner classrooms, and many teachers believe that they make words easier to remember. The research evidence is thinner than that belief. Reviews have found few well-designed studies (Davis, 2017; Hamilton et al., 2024), and a meta-analysis found that the benefit of songs for vocabulary is smaller for children than for adults (Murphy Odo, 2022).

One practical reason the question has been hard to study is that teaching a chosen set of words through song requires a song written around those words. Generative artificial intelligence (AI) now allows a teacher without musical training to produce such a song. This makes a direct classroom comparison possible: the same learners, taught some words through a song and others through the flashcards they already use, and tested on whether they can still produce the words weeks later.

1.2 Statement of the Problem

Three limits in existing research leave teachers of young EFL learners in Taiwan without a clear answer. The first concerns what is measured. Gains in recognizing a

© 2026 Farah M. · Portfolio sample for review only · Not for reuse or redistribution · farah.m.academic@gmail.com

Chapter One · Research Questions

within one class. Each learner is taught 10 of the 20 target words through each method, so each learner's scores on the words taught through song-integrated instruction can be compared with the same learner's scores on the words taught through flashcard instruction. The study also asks how the learners themselves experience the two methods.

1.4 Research Questions

The study addresses three research questions (RQs):

RQ1. Is there a significant difference between song-integrated instruction and flashcard instruction in Grade 4 Taiwanese EFL learners' productive knowledge of the target words on the immediate post-test?

RQ2. Is there a significant difference between song-integrated instruction and flashcard instruction in these learners' retention of that productive knowledge 2 weeks and 7 weeks after the immediate post-test?

RQ3. How do these learners perceive song-integrated instruction and flashcard instruction as ways of learning new words?

RQ1 and RQ2 are answered by comparing each learner's scores on the two sets of words with a paired significance test (Section 3.8.3). RQ3 is answered with a learner questionnaire (Section 3.6.5).

1.5 Significance of the Study

The study contributes in three ways. For teachers, it compares two practical ways of teaching vocabulary under the same classroom conditions, and so bears on a choice that teachers in similar after-school programs make regularly: whether a song is worth the class time that flashcards would otherwise use. Both methods are inexpensive, and song-integrated instruction needs only an audio player and a screen.

© 2026 Farah M. · Portfolio sample for review only · Not for reuse or redistribution · farah.m.academic@gmail.com

Course Paper · EFL Students' Use of ChatGPT Feedback

Theory and Composition of Teaching (Prof. Kai-Lin Wu), Spring 2025. A multiple-case study of how three Taiwanese undergraduates accepted, rejected or modified ChatGPT feedback when revising argumentative essays, and whether their rubric-based scores improved. Shown: the introduction, purpose and research questions.

Introduction

Artificial-intelligence writing assistants have begun to reshape how second-language writers approach revision. In particular, ChatGPT can provide instant suggestions on coherence, word choice, and organization—feedback that traditionally depends on a teacher's limited time. Recent work by Su et al. (2023) argues that ChatGPT can help EFL students improve their reasoning structure. However, there are also questions about how students really interact with the tool's instruction and if such interactions result in quantitative progress, while some students adopt AI suggestions verbatim, others discard them, and a third group adapts the feedback to fit their own rhetorical goals. Understanding these different patterns of use is essential if teachers hope to integrate AI meaningfully rather than superficially. Guided by this need, the present case study follows three Taiwanese undergraduates as they revise argumentative essays with ChatGPT's help. By combining a behavioral analysis of their revision moves with rubric-based scoring of their drafts, the study seeks to clarify both the nature and the impact of AI-mediated feedback in an EFL writing context.

Purpose of the study

This case study examines how EFL students interact with ChatGPT feedback in an authentic course context and whether that engagement leads to demonstrable improvements in writing quality. Specifically, the study revisits the potential of ChatGPT identified by Su et al. (2023)

© 2026 Farah M. · Portfolio sample for review only · Not for reuse or redistribution · farah.m.academic@gmail.com

Course Paper · Research Questions and Methods

and extends Koltovskaia et al.'s (2024) engagement framework by tracing three students' acceptance, rejection, or modification of AI suggestions and linking those revision choices to changes in rubric-based writing scores.

Research questions

Guided by the work of Su et al. (2023) and Koltovskaia et al. (2024), this study addresses the following research questions:

1. How does the EFL student behaviorally engage with ChatGPT feedback to revise argumentative essays?

2. Does ChatGPT feedback help students improve their argumentative writing performance?

Methods

Data collection:

The participants were three Taiwanese undergraduates enrolled in a university-level EFL writing course at Tunghai University. Each student composed two versions of an argumentative essay in response to the prompt "Is social media beneficial for college students?" The first draft was produced independently, with no external input. Upon completion, each student submitted this draft to ChatGPT—using system prompts adapted from Su et al. (2023)—and reviewed the AI's revision suggestions. They then incorporated those suggestions to generate a second draft. For each participant, the following data were compiled for analysis: (a) the complete text of ChatGPT's feedback on Draft 1; (b) both essay

© 2026 Farah M. · Portfolio sample for review only · Not for reuse or redistribution · farah.m.academic@gmail.com

Full papers available on request

The complete thesis proposal includes the literature review, the pilot study, the main-study design, data analysis procedures, ethical considerations and appendices. The complete case study (17 pages) includes the findings for each participant, the scoring rubric and inter-rater results.

Farah M.

farah.m.academic@gmail.com

M.A. in Teaching English as a Foreign Language, Tunghai University, Taichung, Taiwan

© 2026 Farah M. · Portfolio sample for review only · Not for reuse or redistribution · farah.m.academic@gmail.com