http://e-flt.nus.edu.sg/

Electronic Journal of Foreign Language Teaching
2014, Vol. 11, No. 2, pp. 255–269

Three Types of Musical Instruction: Effects on Young Taiwanese EFL Learners' Word Decoding and Rhyme Production

Feng-lan Kuo
(laflkuo@cc.ncue.edu.tw)
National Changhua University of Education, Taiwan

Yihsiang Kuo
(annakuo168@gmail.com)
National Defense University, Taiwan

Yen-hsin Chen
(skinny116@yahoo.com.tw)
Wenhua Elementary School, Taiwan

Robert Pierce
(bobpierce57@gmail.com)
National Changhua University of Education, Taiwan

Abstract

This study compared effectiveness of song versus chant instruction against a combination of these on word decoding (word reading or sounding out words) and rhyme production among four groups of Taiwanese EFL fourth graders. Three intact classes were randomly assigned to receive one of three musical instruction types: song only, chant only, or combination of song and chant instruction, whereas a fourth (control) group received its regular instruction. The instructional period lasted for five weeks with eighty minutes of instruction per week. Researcher-developed sight word decoding and rhyme production tests rated each type of musical instruction. Decoding results for sight words and all words (sight words plus non-words) showed that: (a) three experimental groups (Song, Chant, S + C) not only made significant progress but also remarkably outperformed the Control Group on the decoding posttest; (b) the S + C Group significantly outperformed the Chant Group on decoding posttest; (c) non-significant differences emerged between the Chant Group and other experimental groups. Rhyme production posttest scores regarding real words and all words (real words plus non-words) indicated: (a) all groups significantly progressed in rhyme production; (b) three experimental groups significantly excelled the Control Group; (c) non-significant differences arose among experimental groups. ANOVA results of non-word scores on both decoding and rhyme production posttests revealed non-significant differences among four groups. Ranking of four groups' gain scores (posttest minus pretest) across decoding and rhyme production is consistent: S + C > Song > Chant > Control. Four educational implications and several suggestions for future research are provided based on results of this study.

1 Introduction

Reading plays a significant role in early literacy development. Snow, Burns, and Griffin (1998) stated that children who read well in the early grades can get off to a right start, while those who have lower reading proficiency often stay behind when it comes to academic achievement. In addition, the National Reading Panel (2000) identified phonemic awareness, alphabetic principle, text fluency, vocabulary and comprehension as the five elements of early literacy acquisition for beginning readers. Among these elements, phonemic awareness (PA) and alphabetical principle are the two fundamental factors conditioning early reading success. PA is the ability to manipulate the sounds (including syllables, onsets, rhymes, and phonemes) in spoken words and to understand that spoken words and syllables are made up of sequences of speech sounds (Yopp, 1992). In contrast, the alphabetic understanding refers to being aware that words are composed of letters that represent sounds and using systematic relationships between letters and phonemes (letter-sound correspondence) to retrieve the pronunciation of an unknown printed string (decoding) or to spell words (Snow et al., 1998). The National Reading Panel (2000) further found that phonemic awareness instruction helped students learn to decode unfamiliar words. Likewise, PA training is known to improve learners' word and pseudo word reading in the process of learning to read, which means children can learn to decode novel words by blending and can remember familiar words sharing similar spelling patterns by segmenting (Ehri & Roberts, 2006). Earlier studies conducted in the L1 context also stressed the importance of word reading (Adams, 1990; Ehri, 2005; Ehri & McCormick, 1998; Ehri & Saltmarsh, 1995; Helman & Burns, 2008). The close connection between phonemic awareness and alphabetical principle is thus established. On the other hand, the process of learning sight words enables learners to decode, analogize, predict and produce unfamiliar words by sight (Ehri, 2005).

However, one of the biggest challenges of PA training is that discrete sounds are without boundaries in speech to mark phonemes within words. A more accessible aspect of PA lies in skills of rhyme awareness (including rhyme recognition and production). Rhymes are more salient units and more predictable in pronunciation than vowels alone. In an L1 context, Bryant, Bradley, MacLean, and Crossland's (1989) longitudinal study confirmed that nursery rhymes (a type of commonly known children's music) have positive effects on developing kindergarteners' phonological skills, especially in rhyme and phonemes. For EFL beginners, discriminating onsets and rhymes in nursery rhymes (also called songs here) pave an easier route to decoding (Adams, 1990).

On the other hand, chants resembling songs in content, rhythm, tempo, stress, and intonation except lacking precise melodic lines, can reinforce grammar and pronunciation (Graham, 2006). However, previous L1 studies used either songs as teaching material (Bryant, MacLean, Bradley, & Crossland, 1990; MacLean, Bryant & Bradley, 1987) or a combination of song and chant (song + chant) instruction (Baker, Fernandez-Fein, Scher, & Williams, 1998; Fernandez-Fein & Baker 1997) on enhancing learners' word decoding (i.e. word reading or sounding out words) or rhyme awareness (rhyme recognition and/or rhyme production). No comparative study has been conducted to compare effects of the three types of instruction (song, chant, song + chant) on learners' word decoding and rhyme production. Nevertheless, to better facilitate EFL learning, research interest targeted questions about which one among the three types of musical instruction would work (significantly) better or the best in terms of word decoding and/or rhyme production. This comparison study was thus designed and conducted in the academic fall semester of 2012 to explore this question.

Furthermore, regarding tasks used to measure learners' rhyme awareness, previous L1 studies adopted both rhyme recognition and rhyme production tasks (Fernandez-Fein & Baker, 1997; MacLean et al., 1987) or only rhyme recognition (Bryant et al., 1990), or composite scores of rhyme recognition and rhyme production (Foy & Mann, 2001). Nevertheless, in the Taiwanese EFL context, assessing learners' rhyme awareness performance mostly involved implementation of rhyme recognition tasks (Kuo, 2012; Lin, 2010; Liu, 2005; Yen, 2004). Rhyme production tasks thus were often neglected. To fill such a research gap and identify a more or the most effective approach among the three types of musical instruction, the research calls for an empirical study

comparing effects of three types of musical instruction on Taiwanese EFL children's rhyme production as well as word decoding. This study was thus conducted to fill such a gap. The research questions were posed as follows:

1. Can each of three types of instruction (song, chant, and song + chant) significantly enhance Taiwanese EFL fourth graders' word decoding?
2. Is there a significant difference among the three types of instruction in terms of effect on Taiwanese EFL fourth graders' word decoding?
3. Can each of the three types of instruction significantly enhance Taiwanese EFL fourth graders' rhyme production?
4. Is there a significant difference among the three types of instruction in terms of effect on Taiwanese EFL fourth graders' rhyme production?

2 Literature review

2.1 Effectiveness of three types of instruction: Song, chant, and song + chant

Theoretically, music creates the "song-stuck-in-my-head" phenomenon according to Murphey (1992). It roots the content firmly in the mind through involuntary rehearsal and recall, which is vital to language learning. Empirical studies have been scarce in showing the effectiveness of music in language acquisition. Among various types of music, songs have proven effective in enhancing sight word decoding and rhyme awareness (Adams, 1990; Goswami, 2001). In an L1 context, Baker et al. (1998) found kindergartners' knowledge of nursery rhymes as the strongest predictor of word attack (non-word reading) and word identification skills for second graders, accounting for 36% and 48% of variance, respectively. Bryant et al.'s (1990) study including a pathway analysis showed a route from nursery rhymes to rhyme awareness and an independent route from nursery rhymes to phonemic awareness to reading. Based on Bryant et al.'s study, Goswami (2001) further proposed that songs can promote word reading, since consistent spelling sequences are prominent in nursery rhymes (e.g. *light*, *fight*, and *night*) and even beginners can make analogies between shared spelling in word patterns. This indicates that songs not only facilitate word reading but rhyme awareness.

In the Taiwanese EFL context, various types of music have been used to promote different components of language learning. For instance, songs and chants were used to facilitate young children's rhyme recognition in Yen's (2004) study. Chen (2010) adopted pop songs to enhance sight word acquisition and oral reading fluency. Tseng (2007) and Huang (2010) showed the effectiveness of chants alone on elementary school students' oral reading fluency. Nevertheless, chants resembling songs except lacking precise melodic lines are also useful materials for practicing sight words and rhyme awareness (Abbott, 2000). Since no comparative studies have been conducted to examine respective effects of songs, chants, and the combination of songs + chants on word decoding and rhyme production, this topic warrants exploration.

2.2 Measurement of word decoding

In this study, frequently used standardized word reading tests for elementary school students, including the *Test of Word Reading Efficiency* (TOWRE) and *Dynamic Indicator of Basic Early Literacy Skills* (DIBELS) were reviewed. TOWRE (Torgesen, Wagner, & Rashotte, 1999) has been used to assess learners' competence to decode words accurately and fluently within 45 seconds for each subtest. TOWRE contains two parts, including the *Sight Word Efficiency* (SWE) and the *Phonemic Decoding Efficiency* (PDE). The SWE incorporates the real printed words, while the PDE contains pronounceable printed non-words. The primary selection of sight words is based on the Fry 1000 Instant Word List (Fry, Kress, & Fountoukidis, 2004). Likewise, DIBELS is a fluency measurement, employed to test students from kindergarten to sixth grade. The *DIBELS Nonsense Word Fluency* (DIBELS NWF) is a standardized measurement of alphabetic principle, reflecting commonly used letter-sound correspondence, and is a prerequisite for word identifica-

tion (Kaminski & Good, 1996). Examinees are presented a sheet of paper containing randomly ordered VC and CVC non-words (e.g. *ov*, *sig*), then asked to produce the individual letter sound of each word or read the whole non-word within one minute.

PDE in TOWRE and DIBELS NWF are both standardized tests and have good reliability and validity. The time to administer them is quite short and test administration is easy. A pronunciation guide is provided to score non-word production. Nevertheless, PDE is relatively easier to score since its instruction is simpler and the correct pronunciation is provided in real word samples compared to DIBELS NWF. Most importantly, TOWRE includes both PDE and SWE, while DIBELS NWF only tests non-word decoding ability. Therefore, TOWRE was adopted in the present study to design the word decoding tests.

2.3 Measurement of rhyme awareness

Previous L1 studies categorized measurement of rhyme awareness as rhyme recognition (Bradley & Bryant, 1983; Lenel & Cantor, 1981; Read, 1978), rhyme production (MacLean et al., 1987), or combination of both tasks (Foy & Mann, 2001). Choice of rhyme awareness tasks is determined by research purpose and characteristics of participants. For measuring younger children's rhyme recognition, visual realia, such as pictures or word cards, are often provided to lessen cognition load: for example, Bradley and Bryant's (1983) rhyme oddity tasks, and Lenel and Cantor's (1981) forced-choice rhyme recognition task. Nevertheless, difficulty of rhyme recognition tests may be affected by formal similarity of stimulus words and non-rhyming choices (Foy & Mann, 2001).

By contrast, rhyme production tasks (MacLean et al., 1987) can test real ability in terms of grasp and production of rhyming words, even non-words acceptable. Foy and Mann (2001) adopted composite scores of both tasks as rhyme awareness performance, yet type of rhyme recognition was not specified. Additionally, in prior studies conducted in the Taiwanese EFL context measuring learners' rhyme awareness (Kuo, 2012; Lin, 2010; Liu, 2005; Yen, 2004), tasks were mostly rhyme recognition. Conversely, this study targets rhyme production tasks.

3 Method

3.1 Participants

This experiment lasted seven weeks (from September to October, 2012), including five weeks of instruction in addition to one week apiece for pretest and posttest. Three months before the implementation of the teaching experiment, the third researcher first contacted the Director of Academic Affairs of a local public school in central Taiwan to inform the research objective of the current study. After receiving his approval, fourth graders from four intact classes were then approached via a parent consent letter and volunteered to be involved in this study. All participants, aging from 10 to 11 years old, started formal English education from the third grade: two periods of English per week in both third and fourth grades. Three classes were randomly assigned as experimental groups: song instruction ($n_1 = 26$), chant instruction ($n_2 = 27$) and combination instruction of song and chant ($n_3 = 30$). The fourth class was the control group receiving no song or chant treatment ($n_4 = 29$). Table 1 below briefly reports numerical and gender data for each group.

Table 1. Number and gender descriptions of each group

	Song Group ($n_1 = 26$)		Chant Group ($n_2 = 27$)		Song + Chant ($n_3 = 30$)		Control Group ($n_4 = 29$)	
Gender	Male	Female	Male	Female	Male	Female	Male	Female
Number	14	12	12	15	18	12	14	15

3.2 Teaching materials

Five researcher-developed songs and chants including sight words and rhyming patterns suitable for the level of participants were designed as teaching materials of the present study. Song melodies were all adopted from the "Sing Along" book series (Lard & Lee, 2008), containing clear beats that can be turned into chant format easily. However, lyrics incorporated into the present study were researcher-developed.

Word selection was divided into three phases. First, referring to monosyllabic words on the subtest – SWE of TOWRE, the list of words with the level of Grade Four to Grade Five in the Vocabulary Quotient (National Taiwan Normal University 2006), and high frequency words occurring in the participants' English textbook of "Longman New Go SuperKids," the research team selected words with the principle of *i + 1* (Krashen, 1981). Second, to combine sight words of target rhyming patterns, 37 word families selected from Wylie and Durrell (1970) and 38 word families from Fry (1998) were employed as references to select the sight words. Finally, five patterns of word families were decided, including *-ine, -y, -ack, -ow,* and *-op*. In the end, a total number of 20 sight words was compiled (see Appendix A). Five irregular sight words – words having uncommon grapheme-phoneme correspondences or spellings – including *here, her, some, every,* and *give* – were selected based on the Dolch Sight Word List (Dolch, 1948) and the Vocabulary Quotient. Frequency of word occurrence in teaching materials followed Blevins' (1998) and Reitsma's (1983) threshold frequency of four times.

3.3 Instruments

The researcher-developed word decoding test and the researcher-developed rhyme production test were employed as instruments. The researcher-developed word-decoding test aimed to assess participants' ability to decode (sound out) sight words and non-words of the same rhyming patterns (see Appendix B). The researcher-developed word-decoding test consisted of 25 items. Selection of sight words in the test followed the phases of creating teaching materials, except addition of five more regular words, and five pronounceable non-words. One point was given for each correct response. The maximum score of the researcher-developed word-decoding test was 25 points. Reliability coefficients (Cronbach alpha) of the researcher-developed sight word decoding test and non-word decoding test were .955 and .862, respectively.

Additionally, the rhyme production test assessed participants' rhyme awareness by the number of rhyming words produced within 30 seconds. This test consisted of five questions, each containing examples of rhyming words. For instance, the words *fat* and *hat* rhyme with "*cat*". The researcher would ask children to say "what words rhyme with _____?" (Foy & Mann, 2001). The questions include five words selected from sight words of target rhyming patterns: *by, fine, top, back, low* (see Appendix C). Participants were allowed 30 seconds for each question to say rhyming words. Since more than one word might be produced, one point would be given for each rhyming word. Non-words with target rhyming patterns were also counted as correct responses. The reliability coefficient (Cronbach alpha) of the researcher-developed rhyme production test was .676.

3.4 Research design and procedures

The researcher-developed word decoding and rhyme production tests served as pretests and posttests. The experiment of the present study lasted seven weeks, including five weeks of instruction in addition to one week apiece for pretest and posttest. During intervention, each experimental group received either song, chant, or combined song-chant instruction for two forty-minute sessions (80 minutes) per week. The song instruction group focused on familiarizing participants with word decoding and rhyme production by stressing melodies, while chant instruction group emphasized beat or rhythm. The combined song-chant instruction group mixed melodies and beats but

devoted half the time to song instruction and half to chant instruction. The control group received only pretest, posttest and regular instruction without song or chant treatment.

3.5 Data collection and analysis

Both researcher-developed word decoding and rhyme production tests served as pre- and post-tests to assess participants' word decoding and rhyme production before and after the experiment. Paired-samples *t* tests ascertained whether each type of instruction (song, chant, song + chant) can significantly improve participants' word decoding and rhyme production, respectively (Research Questions 1 and 3). One-way ANCOVA was implemented to accommodate significant differences on pretests and explore significant differences (if any) among three types of instruction regarding effects on word decoding (Research Question 2) and rhyme production (Research Question 4).

4 Results and discussion

4.1 Effects of three types of musical instruction on word decoding

Statistical results of paired-samples *t* tests in Table 2 suggest three types of musical instruction (song, chant, song + chant) significantly enhanced overall word, sight word, and non-word decoding. This finding concurs with prior studies claiming that songs (Bryant et al., 1990; MacLean et al., 1987; Murphey, 1992; Wallace, 1994), chants (Huang, 2010; Tseng, 2007), or songs plus chants (Baker et al., 1998; Fernandez-Fein & Baker, 1997) could assist language learners' word memorization and/or text recalling due to repetitions, melodies and/or rhythmic beats, which in turn remarkably boosts word decoding or oral reading fluency.

Results of the one-way ANOVA in Table 2 were significant on pretest scores regarding overall word decoding and sight word decoding. Such significant results suggest one-way ANCOVA should be used to adjust marked differences in pretest scores among four groups when analyzing posttest scores. Table 3 summarizes both one-way ANCOVA results of posttest scores on overall word decoding as well as sight word decoding and one-way ANOVA results of posttest scores on non-word decoding.

Table 2. Results of paired-samples *t* tests and ANOVAs on four groups' pretest, posttest, and gain scores for overall word, sight word and non-word decoding

1. Overall word decoding (25 words)									
Group	Song (n_1 = 26)		Chant (n_2 = 27)		S + C (n_3 = 30)		Control (n_4 = 29)		
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	p_2
Pretest	9.00	8.50	9.48	9.52	8.83	8.26	13.72	8.69	*.000
Posttest	13.58	10.32	12.11	10.25	14.83	8.58	14.48	8.58	*.000
Gain	4.58	4.82	2.63	3.16	6.00	4.91	0.76	3.00	*.000
p_1	*.000		*.000		*.000		.184		
2. Sight word decoding (20 words)									
Group	Song (n_1 = 26)		Chant (n_2 = 27)		S + C (n_3 = 30)		Control (n_4 = 29)		
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	p_2
Pretest	7.27	6.74	7.78	7.66	7.72	6.61	11.00	6.83	*.000
Posttest	10.77	8.23	9.74	8.31	12.07	6.89	11.34	6.93	*.000
Gain	3.50	3.68	1.96	2.89	4.80	3.82	0.34	2.42	*.000
p_1	*.000		*.002		*.000		0.450		
3. Non-word decoding (5 words)									
Group	Song (n_1 = 26)		Chant (n_2 = 27)		S + C (n_3 = 30)		Control (n_4 = 29)		
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	p_2
Pretest	1.73	1.87	1.70	1.92	1.57	1.81	2.72	.36	.641
Posttest	2.81	2.15	2.37	2.10	2.77	1.87	3.14	1.88	.560
Gain	1.08	1.32	0.67	0.92	1.20	1.42	0.41	1.18	.062
p_1	*.000		*.001		*.000		0.070		

Note: p_1 = *p* value of paired-samples *t* tests; p_2 = *p* value of ANOVAs; Word Decoding Pretest maximal scores = 25. Sight Word Decoding Pretest maximal scores = 20. Non-word Decoding Pretest maximal scores = 5.

Ranking of these groups' adjusted mean scores on word decoding posttest in Table 3 is consistent across overall word and sight word: Song + Chant > Song > Chant > Control. Results of one-way ANCOVA on posttest scores for overall word decoding and sight word decoding were significant ($p < .05$), suggesting marked intergroup differences in scores of overall word decoding and sight word decoding on the posttest. Bonferroni post hoc comparisons were conducted to assess pairwise differences among four groups in terms of overall word decoding and sight word decoding.

Table 3. Summary of ANCOVAs on word decoding posttest scores by word decoding pretest scores for all words and sight words and ANOVA on posttest scores on non-word decoding

	Overall word decoding (25 words)		Sight word decoding (20 words)		Non-word decoding (5 words)	
	Mean	*Adjusted. M.*	*Mean*	*Adjusted. M.*		*Mean*
Song	13.58	14.84	10.77	11.84	Song	2.81
Chant	12.11	12.90	9.74	10.31	Chant	2.37
S + C	14.83	16.28	12.07	13.14	S + C	2.77
Control	14.84	11.15	11.34	8.75	Control	3.14
p_1		*.000		*.000	p_2	.560

Note: p_1 = *p* value of one-way ANCOVAs; p_2 = *p* value of one-way ANOVA

Table 4, listing results of post hoc tests, shows the S + C Group significantly outperformed both Chant and Control Groups on the posttest in terms of overall and sight word decoding. The Song Group also significantly outdid the Control Group in both areas. Non-significant differences emerged between the Song and Chant Groups, Chant and Control Groups, and S + C and Song Groups across overall and sight word decoding.

Table 4. Bonferroni post hoc comparisons of posttest scores for word decoding total and sight word decoding among four groups

		Overall word decoding		Sight word decoding	
		M. D.	*p*	*M. D.*	*p*
S + C	Chant	*3.353	*.003	*2.829	*.001
	Song	1.418	.198	1.300	.140
	Control	*5.107	*.000	*4.394	*.000
Chant	Song	-1.934	.088	-1.529	.091
	Control	1.754	.117	1.566	.080
Song	Control	*3.688	*.001	*3.095	*.001

Although results of pairwise comparisons between Song and Chant Groups failed to attain significance, the Song Group consistently outnumbered the Chant Group to some extent. Moreover, the Song Group significantly excelled the Control Group on the posttest in overall and sight word decoding, while the Chant Group did not. Pairwise comparisons imply that song instruction might work better than chant instruction to some extent in enhancing EFL children's word decoding ability. One possible reason for superiority of song instruction might be that the melody in song is more effective than rhythm in chant in terms of text recalling (Wallace, 1994) and/or word memorization, which in turn facilitates word decoding.

Despite the Chant Group failing to surpass the Control Group conclusively on the word decoding posttest, significant results of paired-samples *t* tests for the Chant Group versus non-significant results for the Control Group imply that chant instruction can still effectively enhance EFL children's word decoding abilities. Furthermore, results of ANCOVA show that the S + C Group consistently exceeded the other three groups across two types (overall and sight word) of word decoding scores, which suggests combination of song and chant instruction as definitely better than either chant or song instruction only. Similar ANCOVA results also occurred to two types (all and real word) of rhyme production scores soon reported in the next subsection. One possible reason for the superior effects of song plus chant instruction on word decoding and rhyme production might be that such instruction contains advantages of both song melody and chant rhythm. Another possible reason for such results might be that combination of song and chant instruction provides task variation, which has been proved significantly more effective than using solely constant task in the field of physical education (Weber & Thorpe, 2010) or in the field of language education for learners with autism (Dawson, 1989). The superiority of instruction with task variation might result from the hypothesis that task variation is likely to arouse more positive learning motivation for children. However, it calls for further research to verify this hypothesis. Based on better effects of song + chant instruction on the participants' word decoding and rhyme production than either song or chant instruction in the current study, EFL teachers are strongly urged to integrate songs and chants into instruction to teach sight words or to enhance young learners' word decoding abilities, rather than only using one of them.

Although ranking of gain scores (posttest minus pretest) across three types of decoding scores (overall word, sight word, and non-word) in Table 2 is consistent (Song + Chant > Song > Chant > Control), ANOVA results of posttest non-word decoding in Table 3 did not attain significance ($p < .05$), suggesting non-significant intergroup differences in decoding among four groups. One possible reason for such non-significant ANOVA results might be that the number of non-word testing items was too small and thus it is much more difficult to reach a significant level. With double

or triple the number of non-word testing items, ANOVA results of non-word decoding scores on the posttest might attain significance. Another reason for the present results is that the Control Group manifested the highest overall word decoding scores on the pretest before the experiment; five weeks of song and/or chant instruction might not be long enough for experimental groups to apply decoding rules and knowledge to non-words or new words not heretofore encountered.

4.2 Effects of three types of instruction on rhyme production

Table 5 displays results of paired-samples *t* tests and one-way ANOVAs on four groups' pretest, posttest, and gain scores for rhyme production of real words, non-words, and all words (real words plus non-words). Results of paired-samples *t* tests indicate four groups progressing significantly on rhyme production in terms of real words, non-words, and all words. Such significant results across four groups and three types of rhyme production scores partially differ from decoding results of paired-samples *t* tests in Table 2, non-significant for the Control Group across three types of decoding scores (overall, sight words, and non-words). One possible reason for such significant results of the paired-samples *t* tests on rhyme production scores across four groups and three types of scores might be that child language learners are more sensitive to rhyming rules, making it easier for them to acquire rhyming rules or awareness and apply them to non-words or new words never encountered.

Table 5. Results of paired-samples *t* tests and ANOVAs on four groups' pretest, posttest, and gain-scores for rhyme production of real words, non-words, and all words

1. All words (pre-range = 0-8; post-range = 0-23)									
Group	Song (n_1 = 26)		Chant (n_2 = 27)		S + C (n_3 = 30)		Control (n_4 = 29)		
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	p_2
Pretest	0.88	1.61	0.74	1.20	1.03	1.96	1.55	1.62	*.000
Posttest	6.58	5.68	5.07	5.11	7.13	5.74	2.83	2.89	*.000
Gain	5.70	4.76	4.33	4.62	6.10	4.88	1.31	2.36	*.000
p_1	*.000		*.000		*.000		*.006		
2. Real words (pre-range = 0-7; post-range = 0-15)									
Group	Song (n_1 = 26)		Chant (n_2 = 27)		S + C (n_3 = 30)		Control (n_4 = 29)		
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	p_2
Pretest	0.73	1.31	0.67	1.14	0.97	1.77	1.41	1.52	*.000
Posttest	4.62	3.95	3.63	3.60	5.50	4.58	2.21	2.34	*.000
Gain	3.89	3.23	2.97	3.29	4.53	3.53	0.80	1.88	*.000
$\boldsymbol{p}_1$	*.000		*.000		*.000		*.031		
3. Non-words (pre-range = 0-2; post-range = 0-10)									
Group	Song (n_1 = 26)		Chant (n_2 = 27)		S + C (n_3 = 30)		Control (n_4 = 29)		
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	p_2
Pretest	0.15	.46	0.07	.27	0.06	.25	0.14	.44	.706
Posttest	1.96	2.60	1.44	2.38	1.63	2.13	0.66	.86	.239
Gain	1.81	2.42	1.37	2.39	1.57	2.19	0.52	.95	.105
p_1	*.001		*.006		*.001		*.007		

Note: p_1 = *p* value of paired-samples *t* tests; p_2 = *p* value of ANOVAs; pre-range = range of pretest scores; post-range = range of posttest scores

This reason can also logically account for non-significant ANOVA results of non-word rhyme production scores on the posttest, suggesting non-significant differences among groups in non-word rhyme production after musical treatments. With acquiring rhyming rules or awareness being

easier for EFL youngsters, the Control Group without musical treatment made as significant progress in non-word rhyme production as the three experimental groups did, leading to non-significant differences in non-word rhyme production scores on the posttest across groups.

To further investigate if there is a significant difference in gain scores (posttest minus pretest) among four groups, one-way ANOVA was conducted. Results of the one-way ANOVA of the rhyme production pretest scores are significant in terms of real words and all words, suggesting one-way

ANCOVA should be used to adjust marked intergroup mean differences in pretest scores when analyzing rhyme production posttest scores. Table 6 depicts results of the one-way ANCOVA on rhyme production posttest scores by rhyme production pretest scores regarding real words and all words as well as results of one-way ANOVA on posttest scores on non-word rhyme production.

Table 6. Results of one-way ANCOVAs on thyme production posttest scores by rhyme production pretest scores for all and real words and one-way ANOVA on posttest scores on non-word rhyme production

	All words (range: 0-23)			Real words (range 0-15)			Non-words (range: 0-10)		
	Mean	SD	Adjusted mean	Mean	SD	Adjusted mean		Mean	SD
Song	6.58	5.68	6.89	4.62	3.95	4.96	Song	1.96	2.60
Chant	5.07	5.11	5.64	3.63	3.60	4.07	Chant	1.44	2.38
S + C	7.13	5.74	7.19	5.43	4.55	5.42	S + C	1.63	2.13
Control	2.83	2.89	1.97	2.21	2.34	1.50	Control	0.66	0.86
Total	5.39	5.20		3.97	3.85		Total	1.41	2.11
p_1			*.000			*.000	p_2	.118	

Note: p_1 = p value of one-way ANCOVAs; p_2 = p value of one-way ANOVA

Like results of word decoding posttest scores, ranking of the four groups' adjusted mean scores on the rhyme production posttest is consistent across two types (all words and real words) of scores: Song + Chant > Song > Chant > Control. Also similar to results of one-way ANCOVA on word decoding posttest scores, results of the one-way ANCOVA on rhyme production posttest scores for all words and real words were significant ($p < .05$), suggesting noteworthy differences among four groups in posttest scores for rhyme production of all words and sight words. Bonferroni post hoc comparisons assessed pairwise differences among four groups in rhyme production of all words and real words. Table 7 presenting results of post hoc tests shows that all three experimental groups (S + C, Song, and Chant) significantly outnumbered the Control Group on posttest in terms of all words and real words, yet non-significant differences arose among the three experimental groups. Such results suggest three types of instruction can significantly increase EFL children's rhyme production ability.

Table 7. Bonferroni post hoc comparisons of posttest scores for rhyme production of all words and sight word among four groups

		All words (range: 0-23)		Sight words (range 0-15)	
		M. D.	*p*	*M. D.*	*p*
S + C	Chant	1.546	.159	1.344	.090
	Song	.295	.788	.457	.565
	Control	*5.215	*.000	*3.911	*.000
Chant	Song	-1.250	.269	-.888	.276
	Control	*3.670	*.001	*2.567	*.002
Song	Control	*4.920	*.000	*3.454*	*.000

In spite of non-significant differences among three experimental groups in rhyme production posttest scores and gain scores (posttest minus pretest) in terms of all words and real words, the S

+ C Group consistently surpassed Song Group and Chant Group in sequence. Such results imply that to some extent song plus chant instruction might be more effective in boosting EFL children's rhyme production ability than song instruction or chant instruction. These rhyming gain score results are in line with the decoding results that the S + C Group obtained the highest gain scores, sequentially followed by the Song Group, Chant Group, and Control Group. Two possible reasons (containing advantages of both song melody and chant rhythm as well as task variation) for the superior effects of song plus chant instruction on word decoding and rhyme production were discussed in the previous subsection on page 262.

5 Conclusions, educational implications, and limitations

This study compared effects of three types of musical instruction (song, chant, song + chant) on Taiwanese EFL children's word decoding and rhyme production. Three conclusions emerge from statistical analyses of decoding and rhyme production data. First, the three types of instruction (song, chant, song plus chant) could significantly bolster EFL children's sight word decoding and rhyme production of real words with song plus chant instruction best, sequentially followed by song instruction and chant instruction. A possible reason for this superiority of song plus chant instruction might be that such a combination instruction holds advantages of both song melody and chant rhythm. Second, effects of song seem to be *relatively* better than those of chant in advancing EFL young learners' decoding and rhyme production. One possible reason is that melodies in songs may be more effective than rhythmic beats in chants in word memorization and text recalling (Wallace, 1994), which in turn facilitate word decoding and rhyme production. Third, child language learners may be more sensitive to rhyming rules, accelerating their acquisition of rhyming rules or awareness faster than that of decoding knowledge or skill. This conclusion is drawn based on significant results of paired-samples *t* tests across three types of rhyme production scores for the Control Group versus non-significant results of paired-samples *t* tests across types of word decoding scores for the Control Group, suggesting rhyme production as easier than word decoding so that the Control Group without musical treatment made notable progress on rhyme production, like all three experimental groups did, while failing to do so on decoding posttest. The result complies with the finding of Helfgott (1976) that partial blending C-VC is easier than phoneme blending of CVC words for L1 kindergartners. The result is also consistent with the finding of Seymour and Evans (1994) that onset-rhyme blending is easier than phoneme blending for L1 5-year-olds and 6-year-olds. Among the 25 words tested in the researcher-developed word decoding test, 10 words are of CV structure while 15 words are of CVC structure. This indicates that 60% of the test items require the children to perform phoneme blending. In contrast, in the rhyme production task the children are only required to do the onset-rhyme blending. Our finding therefore supports the claim of Pufpaff (2009) that phonological sensitivity represents a continuum from awareness of larger units (onset-rhyme) of spoken language progressively to smaller units (phoneme) of spoken language.

Four educational implications for EFL teachers and textbook designers are provided based on the findings of the current study. First, teachers are strongly recommended to adopt both songs and chants to augment EFL children's word decoding and rhyme production. As for the sequence of combining song with chant activities, chants with clear beats helping students familiarize with rhythm of content can be first used to teach rhythmic patterns and then melody in a song can be added to bring variation to facilitate memory of rhythmic patterns to magnify learning effects. Second, as with researcher-developed songs and chants in this study, textbook designers can incorporate high-frequency and/or sight words into popular nursery rhymes or chants to expedite learning in a fun and repetitive way. Third, textbook designers or teachers can add rhyming words sharing similar patterns with sight words in songs or chants to enhance both rhyme awareness and word decoding ability simultaneously. Fourth, since Perfetti and Marron (1998) found evidence that adult illiterates lack phonemic awareness and decoding skills, it is suggested that there are fundamental similarities between adults' and children's literacy development. Therefore, it can be inferred from the results of the present study that discrimination in onsets and rhymes and training in

word decoding using songs and chants can be effectively applied to adult EFL beginners who lack PA and word decoding skills or who need remedial education. Moreover, the findings of this case study can be potentially applied to EFL learners of different language backgrounds but with similar English reading proficiency. Future studies can be conducted to compare/verify the effectiveness of different types of musical instruction on EFL beginners at different ages and/or with different language backgrounds.

Some limitations of this study followed by suggestions for future research require acknowledgment. First, the rhyme production task serving as a measure of learners' rhyme awareness performance was a preliminary attempt in a Taiwanese EFL context, the reliability coefficient of this task was .676, slightly short of the .7 threshold. Nevertheless, the reliability coefficient of the rhyme production task used in previous studies was unknown (Fernandez-Fein & Baker, 1997; Foy & Mann, 2001; MacLean et al., 1987). Second, intervention time should span longer than five weeks for future studies to verify if longer intervention time yields different or better outcomes. Third, future studies can implement delayed posttest to assess retention effects of songs or chants on decoding (real words and non-words) as well as rhyme production (real words and non-words), given adequate assisted resources. Fourth, prior L1 studies gauging effects of songs or chants on rhyme awareness (Fernandez-Fein & Baker, 1997; MacLean et al., 1987) used rhyme recognition and rhyme production tasks simultaneously. In future research, it appears worthwhile to conduct comparative studies utilizing songs and chants to measure both rhyme recognition and production. Fifth, diverse genres of music can be used in future studies to probe effects of each on word decoding and rhyme awareness. Finally, a study is needed to verify if acquisition of rhyme awareness is easier than that of (sight) word decoding for young EFL learners or vice versa and explore possible reasons.

Acknowledgement

The authors immensely appreciate the constructive and insightful suggestions of two anonymous reviewers for improving the draft of this article.

References

Abbott, M. (2000). Using music to promote L2 learning among adult learners. *TESOL Journal, 11*(1), 10–17.

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT press.

Baker, L., Fernandez-Fein, S., Scher, D., & Williams, H. (1998). Home experiences related to the development of word recognition. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 263–287). Hillsdale, NJ: Erlbaum.

Blevins, W. (1998). *Phonics from A to Z: A practical guide*. New York, NY: Scholastic Professional Books.

Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read: A casual connection. *Nature, 301,* 419–421.

Bryant, P. E., Bradley, L., MacLean, M., & Crossland, J. (1989). Nursery rhymes, phonological skills and reading. *Journal of Child Language, 16*(2), 407–428.

Bryant, P. E., MacLean, M., Bradley, L., & Crossland, J. (1990). Rhyme, alliteration, phoneme detection, and learning to read. *Developmental Psychology, 26*, 429–438.

Chen, Y. C. (2010). *A study of enhancing EFL young learners' sight vocabulary acquisition and oral reading fluency through pop song singing-reading instruction* (Unpublished master's thesis). National Cheng Kung University, Tainan, Taiwan.

Dawson, G. (Ed.) (1989). *Autism: Nature, diagnosis, and treatment*. New York, NY: Guilford Press.

Dolch, E. W. (1948). *Problems in reading.* Champaign, IL: The Garrard Press.

Ehri, L. C. (2005). Learning to read words: theory, findings, and issues. *Scientific Study of Reading, 9*(2), 167–188.

Ehri, L. C., & McCormick, S. (1998). Phases of word reading: Implications for instruction with delayed and disabled readers. *Reading and Writing Quarterly, 14*, 135–163.

Ehri, L. C., & Roberts, T. (2006). The roots of learning to read and write: Acquisition of letters and phonemic awareness. In D. K. Dickinson & S. T. Neuman (Eds.), *Handbook of early literacy research* (pp. 113–131). New York, NY: The Guilford Press.

Ehri, L. C., & Saltmarsh, J. (1995). Beginning readers outperform older disabled readers in learning to read words by sight. *Reading and Writing: An Interdisciplinary Journal, 7*, 295–326.

Fernandez-Fein, S., & Baker, L. (1997). Rhyme and alliteration sensitivity and relevant experiences among preschoolers from diverse backgrounds. *Journal of Literacy Research, 29*(3), 433–459.

Foy, J., & Mann, V. (2001). Does strength of phonological representations predict phonological awareness in preschool children. *Applied Psycholinguistics, 22*, 301–325.

Fry, E. (1998). The most common phonograms. *The Reading Teacher*, *51*(7), 620–622.

Fry, E., Kress J., & Fountoukidis, D. (2004). *The reading teacher's book of lists*. Paramus, NJ: Prentice Hall.

Goswami, U. (2001). Early phonological development and the acquisition of literacy. In S. B. Neuman & D. K. Dickson (Eds.), *Handbook of early literacy research* (pp. 111–125). New York, NY: Guilford Press.

Graham, C. (2006). *Creating chants and songs*. Oxford: Oxford University Press.

Helfgott, J. A. (1976). Phoneme segmentation and blending skills of kindergarten children: Implications for beginning reading instruction. *Contemporary Educational Psychology, 1*, 157–169.

Helman, L. A., & Burns, M. K. (2008). What does oral language have to do with it? Helping young English-language learners acquire a sight word vocabulary. *The Reading Teacher, 62*(1), 14–19.

Huang, S. C. (2010). *The effect of reader-theater-jazz chant instruction on elementary school's oral reading fluency* (Unpublished master's thesis). Providence University, Taichung, Taiwan.

Kaminski, R. A., & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review, 25*, 215–227.

Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon Press.

Kuo, M. M. (2012). *The effects of preliminary phonics instruction on EFL fifth graders' oral reading performance* (Unpublished master's thesis). Providence University, Taichung, Taiwan.

Lard, S., & Lee, C. (2008). *Sing along*. Taipei: Ivy League Analytical English.

Lenel, J. C., & Cantor, J. H. (1981). Rhyme recognition and phonemic perception in young children. *Journal of Psycholinguistic Research, 10*(1), 57–67.

Lin, L. Y. (2010). *Effects of using English nursery rhymes on elementary school students' phonological awareness and word recognition skills* (Unpublished master's thesis). National Kaohsiung Normal University, Kaohsiung, Taiwan.

Liu, S. C. (2005). *The study of nursery rhyme teaching to develop children's rhyme awareness* (Unpublished master's thesis). National Taipei University of Education, Taipei, Taiwan.

MacLean, M., Bryant, P. E., & Bradley, L. (1987). Rhymes, nursery rhymes, and reading in early childhood. *Merrill-Palmer Quarterly, 33*, 255–282.

Murphey, T. (1992). *Music & song.* Oxford: Oxford University Press.

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implication for reading instruction.* Washington, DC: Author. Retrieved from http://www.nichd.nih.gov/publications/nrp/smallbook.htm

Perfetti, C. A., & Marron, M. A. (1998). Learning to read: Literacy acquisition by children and adults. In D. A. Wagner (Ed.), *Advances in adult literacy research and development* (pp. 1–43). New York, NY: Hampton Press.

Pufpaff, L. A. (2009). A developmental continuum of phonological sensitivity skills. *Psychology in the-Schools, 46*(7), 679–791.

Read, C. (1978). Children's awareness of sounds, with emphasis on sound system. In A. Sinclair, R. J. Jarvella, & W. J. M. Levelt (Eds.), *The child's conception of language* (pp. 65–82). Berlin: Springer-Verlag.

Reitsma, P. (1983). Printed word learning in beginning readers. *Journal of Experimental Child Psychology, 36*(2), 321–339.

Seymour, P. H. K., & Evans, H. M. (1994). Levels of phonological awareness and learning to read. *Reading and Writing: An Interdisciplinary Journal, 6*, 221–250.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Torgesen, J., Wagner, R., & Rashotte, C. (1999). *TOWRE: Test of word reading efficiency*. Austin, TX: PRD-ED.

Tseng, S. D. (2007). *Effects of teaching of jazz chants on English speaking Performance and attitudes for sixth-graders in southern Taiwan* (Unpublished master's thesis). National Kaohsiung Normal University, Kaohsiung, Taiwan.

Wallace, W. T. (1994). Memory for music: Effect of melody on recall of text. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1471–1485.

Weber, R. C., & Thorpe, J. (2010). Comparison of task variation and constant task methods for severely disabled in physical education. *Adapted Physical Activity Quarterly, 6*(4), 338–353.

Wylie, R. E., & Durrell, D. D. (1970). Teaching vowels through phonograms. *Elementary English, 47*, 787–791.
Yen, T. C. (2004). *Effect of teaching English nursery rhymes and chants on young Taiwanese EFL beginners' rime awareness* (Unpublished master's thesis). Tunghai University, Taichung, Taiwan.
Yopp, H. K. (1992). Developing phonemic awareness in young children. *The Reading Teacher, 45*, 696–703.

Appendices

Appendix A

Teaching materials

1. **Apple Trees**: **-ine**: line, fine, mine (tune: *Sing a Song of Sixpence*) Plant **every** tree of apple All in a line. Water **every** day and they will be fine. When **every** tree all grows tall, Children make a line. **Every** apple on the tree, They will all be mine. **3. A Boy: -ack**: black, back, pack (tune: *The Muffin Man*) Oh, do you know a boy in black with a pack on his back? Oh, yes, I know the boy in black who lives in wonderland. Oh, will you **give** the boy a hand and **give** a toy as a gift? Oh, yes, I will **give** the boy a hand and **give** him what he wants. **4. Four Little Fish: -ow**: show, low, bow (tune: *Six Little Ducks*) Four little fish come for a show. Big ones, small ones swim high and low. But the one little fish with a bow on **her** head, She led the others jumping in **her** show. Jump with **her**! Jump with **her**! All the little fish are jumping to and fro.	**2. Blackbird: -op**: stop, top, hop (tune: *Kookaburra*) Blackbird sits in the old tree top, Looking for **some** places where he could hop. Stop, blackbird! Stop, blackbird! Sing **some** songs for me. Blackbird waits in the starry night, Looking for **some** worms that he could eat. Smile, blackbird! Smile, blackbird! Leave **some** there for me. **5. I Stand: -y**: cry, shy, try (tune: *Head, Shoulders, Knees and Toes*) I stand **here** all alone and cry. I stand **here** all alone and cry. I am shy, but I will try. Hey, everybody please stand up **here**. X2

Appendix B

Researcher-developed Word Decoding Test

Class: _____ Number: _____

No.	Sight Words	Pretest Score	Posttest Score
1	cry		
2	shy		
3	try		
4	by		
5	line		
6	fine		
7	mine		
8	wine		
9	stop		
10	top		
11	hop		
12	shop		
13	back		
14	black		
15	pack		
16	snack		
17	show		
18	low		
19	bow		
20	snow		
	Total score		
	Non-words	**Pretest Score**	**Posttest Score**
1	py		
2	hine		
3	dop		
4	plack		
5	fow		
	Total score		

Appendix C

Researcher-developed Rhyme Production Test

Class: _____ Number: _____

The word cat rhymes with fat and hat.

1. What words rhyme with by?
2. What words rhyme with fine?
3. What words rhyme with top?
4. What words rhyme with back?
5. What words rhyme with low?