UNIVERSITAT DE BARCELONA

M.A. Thesis

***Please don't stop the music*:**
an exploratory study on teaching pronunciation and vocabulary
through songs in different modalities.

Paola Mannarelli

Supervisor: Dr. **Raquel Serrano Serrano**

Applied Linguistics and Language Acquisition in Multilingual Contexts

M.A. Program

Department of Modern Languages and Literatures and
English Studies

Universitat de Barcelona

Academic year: 2018-2019

Facultat de Filologia
Departament Llengües Modernes i Estudis Anglesos

Gran Via de les Corts Catalanes, 585
08007 Barcelona, SPAIN
Tel. +34 934 035 679
Fax +34 933 171 249

**Màster Oficial en Lingüística Aplicada
i Adquisició de Llengües en Contextos Multilingües
LAALCM**

Dr. Raquel Serrano Serrano com a supervisora del treball (Tesina de
(nom i cognoms)

Màster) presentat com a requeriment per a l'avaluació de l'assignatura **Projecte de**

Recerca en Lingüística Aplicada
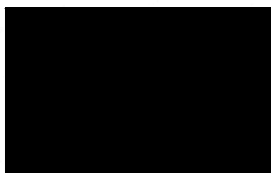
presentat per l'alumne/a: **Paola Mannarelli**
(nom i cognoms)

amb el títol de: *"Please don't stop the music*: an exploratory study on teaching pronunciation and vocabulary through songs in different modalities".

certifico que he llegit el treball i l'aprovo perquè pugui ser presentat per a la seva defensa pública al mes de Septiembre de l'any 2019

I perquè consti i tingui els efectes oportuns signo aquest certificat en

Barcelona, a 6 de setembre de 2019

....

Dr/a. Raquel Serrano

Facultat de Filologia
Departament Llengües Modernes i Estudis Anglesos

Gran Via de les Corts Catalanes, 585
08007 Barcelona, SPAIN
Tel. +34 934 035 679
Fax +34 933 171 249

Official M.A. program in
Applied Linguistics and Language Acquisition in Multilingual Contexts (LAALCM)

Universitat de Barcelona

Non-Plagiarism Statement

This form must be completed, dated and signed and must be included at the beginning of every copy of the M.A. Thesis you submit for assessment.

Name and surnames:	Paola Mannarelli
MA Thesis title:	"*Please don't stop the music*: an exploratory study on teaching pronunciation and vocabulary through songs in different modalities".
Supervisor:	Dr. Raquel Serrano Serrano

I HEREBY DECLARE THAT:

- This MA Thesis that I am submitting for assessment is entirely my own work and I have written it completely by myself.
- I have not previously submitted this work or any version of it for assessment in any other program or institution.
- I have not used any other sources or resources than the ones mentioned.
- I have identified and included the source of all facts, ideas, opinions and viewpoints of others through in-text referencing and the relevant sources are all included in the list of references at the end of my work. Direct quotations from books, journal articles, internet sources or any other source whatsoever are acknowledged and the sources cited are identified in the list of references.

I understand that plagiarism and copying are serious offences. In case of proof that this MA Thesis fails to comply with this declaration, either as negligence or as a deliberate act, I understand that the examiner has the right to exclude me from the assessment act and consequently all research activities conducted for this course will be declared null and the MA Thesis will not be presented for public defence, thus obtaining the lowest qualification.

Date: 6th of September   **Signature:**

Acknowledgments

First and foremost, I would like to thank my research supervisor Dr. Raquel Serrano. Thank you for your patience, valuable suggestions and for not thinking I was crazy.

I sincerely thank Dr. Elsa Tragant and Dr. Joan Carles Mora for recommendations and reference materials given me throughout the course and the thesis process, as I am sure they had better things to do, but helped me anyway.

Special thanks go to my previous students, colleagues and directors, for their unconditional support.

Last but definitely not least, I am so extremely grateful to Pablo for helping me and putting up with me. All. The. Time.

Abstract

Pop songs could provide a noteworthy source of L2 native input both inside and outside the classroom even in settings where the foreign language is not ordinarily spoken. The present study investigates perceptual/receptive and productive learning of English pronunciation and vocabulary through songs in different modalities (with or without lyrics) and instruction (direct or indirect) in the case of teenage Italian learners. Results show that treatment was effective for most tests and they provide support for the inclusion of songs in the L2 class. Overall, Modality did not stand out as a significant variable, while Instruction did only in Vocabulary Production. The study also registered a significant interaction between the no-lyrics modality and the indirect type of instruction, which supports the idea of effective extensive listening, when the activity is not aimed at comprehension and aural input is supported with images.

Table of Contents

1. Introduction

Listening to music is an activity that students engage with outside the classroom, but that can also have a positive impact inside it. For instance, songs help alleviating Foreign Language Classroom Anxiety (FLCA), as confirmed by Dolean (2015) and such feature is essential in lowering the emotional barrier theorized by Stephen Krashen (1982) in his Affective Filter Hypothesis, which might prevent learners from achieving proficiency in an instructed foreign language (FL) setting. Furthermore, research proved that "during an electroencephalogram (EEG), music can change brain waves and make the brain more receptive to learning. Music connects the functions of the right and left hemispheres of the brain so that they can work together and make learning quick and easy." (Rahman, 2007, p.2). Not only that, but songs can also provide "that bit of language that is heard/read and that is slightly ahead of a learner's current state of grammatical knowledge" (Gass & Selinker, 2001, p. 200) which in Krashen's Input Hypothesis (1982) was considered essential for learners to make progress and acquire a language.

Thus, songs have the potential to facilitate a positive FL learning experience and provide for a significant source of second language (L2) input in a context characterized by its lack-there-of. In the present study the students' progress will be analysed both for pronunciation and vocabulary, not only from the point of view of perceptual/receptive[1] skills, but also of productive ones. Moreover, results are going to be scrutinized in order to assess the impact of different input modes (with or without lyrics) and of different instruction types (direct or indirect), thus in the next section, the literature review will discuss relevant studies on songs as significant L2 input and valuable sources for pronunciation and vocabulary teaching, together with studies on different instruction modalities.

---

[1] Following the specific terminology of pronunciation and vocabulary literature, passive skills will be referred to as "perceptual" for pronunciation and "receptive" for vocabulary.

2. Literature Review

2.1 *I Don't Like Mondays (The Boomtown Rats, 1979)*:

"The Motherese of Adolescents"

Teenagers independently spend time listening to music outside the classroom, as pointed out by Miranda (2013), who found that in the US "on average, adolescents listen to music for up to three hours daily and accumulate more than 10,000 hours of active music listening throughout adolescence". That is why pop songs have been defined as "the motherese of adolescents" (Murphey & Alber, 1985, p.1) and such input should also be taken into consideration as valuable source of learning (Schwarz, 2013).

In particular, as far as songs in English are concerned, such reflection can be applied also to contexts where English is not an official language. In Italy, specific setting of this study, *the Observatory on Adolescent Trends and Behaviors* in 2016 carried out a research on around 7,000 Italian teenagers between 13 and 19 and found that 98.5% of them listens to music regularly. Besides, the 2019 Radiomonitor survey emphasised that Italian radios play 51% Italian songs and 49% foreign ones. Due to the recent introduction of the INVALSI (*Istituto nazionale per la valutazione del sistema educativo di istruzione e di formazione*) English national examination, which involves a reading comprehension and a listening comprehension test, though, it becomes even more crucial for Italian students to capitalize on such input.

In Italy, English is the FL that students are more commonly exposed to out-of-class, but the dominance of English is also common in other European areas where there is exposure to more than one FL, like Flanders, where Dutch is the official language of the region, while Dutch, French and German are the official languages of Belgium. Such area was the setting of a study by Peters, Noreillie, Heylen, Bulté and Desmet (2019), which established that English was more present than French in students' out-of-class activities.

The quantity and quality of input are pivotal factors in both first and second language acquisition. However, despite its importance, input has so far been an underestimated element in Second Language Acquisition (SLA) research, mainly due to its variability and the difficulty in measuring both its quantity and its quality (Flege, 2009). In *Give Input a Chance!*, James Flege focused on

reviewing studies in naturalistic settings, where learners are generally exposed to remarkable rates of L2 input even outside of the classroom, but the role of L2 input in FL settings, where the L2 is not used outside the classroom, should deserve just as much attention.

Saito (2019, p.2) highlights that "while immersion/study abroad is commonly conceived as the optimal way to improve such L2 skills, a great number of L2 leaners study their target languages in FL settings. […] SLA in FL settings is limited in terms of both quantity and quality". Due to this lack of FL, learners often have to deal with issues that make comprehension of the aural input difficult (Goh, 1999; Chang & Read, 2006), such as quickness of speech, unknown vocabulary, or failure to connect the spoken and written forms (Chang, 2009) and necessitate strategies to assist listening comprehension. Further difficulties could also arise with learners being concerned about their accented L2 speech (e.g., Derwing, 2003), maybe because they are conscious about native speakers' negative judgement on this respect (Kissling, 2013). Therefore, listening, pronunciation and vocabulary arise as key areas to address students' needs in learning a FL.

2.2 *Learn to Listen (The Ramones, 1989)*:
Listening for pronunciation training

Pronunciation teaching has been overlooked in applied linguistics (Derwing & Munro, 2005). The reasons behind such situation have been studied by Isabelle Darcy (2018) through a survey of teachers' practices and perceptions about pronunciation in the Intensive English Program at Indiana University. The study showed that usually teachers do not include pronunciation training within their lessons due to lack of time, lack of training and lack of importance in assessment. The same discomfort in teaching pronunciation is felt by teachers in Italy (Copland, Garton & Burns, 2014), as such area of instruction is also underdeveloped in the Italian system (Costenaro, 2011).

Pronunciation teaching should encompass a variety of aspects (Goodwin, 2014), which belong to three different areas: segmental (i.e. consonants and vowels), suprasegmental (i.e. word stress and intonation) and fluency. The present study will only focus on segmental features, both from the point of

view of perception and production, due to the shortness of treatment. According to the Speech Learning Model (SLM): "without accurate *perceptual* targets to guide the sensorimonitor learning of FL sounds, production of the L2 sounds will be inaccurate" (Flege 1995, p. 238). This does not imply that all pronunciation production mistakes are due to perceptual ones, but underlines the close connection between the two aspects.

Gilbert (1995) projects these two domains of speech communication in the classroom and reaches the conclusion that listening activities enhance the development of perceptual skills, which then help productive ones improve. Specifically, in relation to listening to songs, the *song-stuck-in-my-head* phenomenon hypothesised by Murphey (1990) as the involuntary mental rehearsal of a song in one's head, could be extremely valuable in order to create or modify existing phonetic categories. Besides, Kissling (2013) and Isaacs (2009) underline the need for pronunciation instruction to be performed via communicative activities, in order to direct the learners' attention to L2 sounds, "perhaps through targeted exposure, focused listening, dictation, transcription, or other means, should be explored and weighed against the potential benefits of explicit phonetics instruction" (Kissling, 2013, p.725).

Listening activities and pronunciation training share the same marginalized role in the FL classroom, where the focus of direct instruction is rarely on such skills. In Beall, Gill-Rosie, Tate, and Matten (2008), it is demonstrated that not enough time is dedicated to listening skills in instructed settings, even though they have been considered crucial even for reading comprehension (Campbell, 2011). In the abovementioned study, Darcy (2018) also suggests that in order for a listening practice to be useful for effective pronunciation training, it would be beneficial to have meaningfully contextualized phonological features with vocabulary items and to increase the amount of times learners are exposed to them. Indeed, listening to songs fulfils both requirements.

Nevertheless with regards to the relationship between music and pronunciation, listening to songs might only be linked to improvements in perceptual skills. In fact, Christiner and Reiterer (2015) in their study "a Mozart is not a Pavarotti" saw that singers outperformed musicians in a foreign accent imitation task, while there was not a significant difference between the former and the latter in terms of perceptual abilities. Thus, the

capability to produce accurate sounds in a FL seems to be more strongly correlated to singing than merely listening to music or playing an instrument, since "singers benefit from heightened vocal motor abilities" (Coumel, Christiner & Reiterer, 2019, p.3).

2.3 *You Don't Learn That In School (Nat King Cole, 1946)*:

Listening for vocabulary teaching

Listening, speaking, reading, and writing are all fundamental skills in learning a language, be it the first or second one, and they all depend on three core aspects: "phonology, vocabulary (lexis), and structure" (Darcy, 2018). As far as vocabulary is concerned, rhythm, one of music components, has been demonstrated to be highly beneficial for rote memory (Medina, 1993), but this is only one of the positive aspects of learning L2 vocabulary through songs. In their study on incidental vocabulary learning by listening to songs in Thai students in Year 6 and 7, Pavia, Webb and Faez (2019) identified six reasons that make songs a remarkable resource to learn an L2:

> First they can provide large quantities of language input. […] Second, corpus-driven studies analyzing pop songs suggest that these songs are repetitive and conversation-like and that the mean speed of speech of 75.49 words per minutes was half the speed of spoken discourse (Murphey, 1990). […] Third […] individuals tend to listen to the same song multiple times. […] Fourth, many of the same words are encountered in different songs. […] Fifth, songs can also assist memory in language acquisition. […] Sixth, it has been demonstrated that the use of songs in FL classrooms tends to decrease anxiety levels in high-anxiety classrooms. (Pavia et al., 2019, p.5)

Despite all these positive aspects, songs are still considered as a one-off treat that teachers give students, rather than a regular activity, as demonstrated by their sporadic presence in textbooks (Tegge, 2015). The same situation can be applied to the research field, where despite an extensive pedagogical literature supporting the efficiency of learning through songs (Tegge, 2015), most studies only focus on participants' self-reports, rather than empirically verified gains (e.g. Pavia et al., 2019; Medina, 1993).

The relationship between vocabulary and listening to songs appears to be bi-univocal, with the latter benefitting the former and vocabulary knowledge enhancing understanding of songs. In fact, Stæhr (2008) found that vocabulary

size could predict 39% of variance in listening scores. In order to build a large vocabulary size, though, Peters et al. (2019) specifically highlighted the importance of out-of-class exposure to the L2, due to time constraints in the classroom.

Nonetheless, listening to songs in the L2 might be linked only to the receptive knowledge of vocabulary and not to the productive one. For example, the already mentioned studies conducted by Pavia et al. (2019) and Medina (1993) only include multiple-choice tests assessing receptive knowledge of target words. According to Nation, though, (1990, p.32) "productive knowledge of a word includes receptive knowledge and extends it" and Schmidt (1990) identifies "recycling" as the key to such extension, which implies that students are presented with the word several times in different contexts. These characteristics are fulfilled by pop-song discourse, as confirmed by Murphey's (1992) corpus-analysis of the top 50 songs in English at that time, where he found that "each word is repeated about three times in an average song of 263 tokens. Actually 25% of the corpus is composed of just 10 different words" (p.771).

2.4 *More than Words (Extreme, 1990)*:
Input modes in SLA instruction

A bimodal input (audio and written or audio and video) is believed to be more beneficial for learning to follow. For instance, Chang (2009), comparing the Reading While Listening (RWL) and the Listening Only (LO) modes from the listening instruction perspective found that reading facilitates listening comprehension. Moreover, understanding what was being said was highly beneficial for the students to keep motivated during the activity. As a matter fact, one of the main issues of LO is its transitory nature, which allows for less control of the input by learners and thus causes greater anxiety and possibly a decrease in motivation (Chang, 2009). Webb and Chang (2012) also established that RWL leads to a higher degree of comprehension than Reading Only (RO), and RWL has especially proved to be useful in terms of linking form to meaning (Webb and Chang, 2012) and of segmenting texts into larger chunks (Brown, Waring & Donkaewbua, 2008).

6

Nevertheless, "it is often hard to make out the words in songs, even for native speakers" (Garnier & Schmitt, 2016, p. 37), so reducing songs to comprehension activities would be rather diminutive. Songs can be presented in different ways, such as LO aural input, bimodal (audio with lyrics or aural input with static images) or multimodal, which includes aural input, lyrics and video or non-static images.

Extensive literature confirms that RWL helps developing listening comprehension skills (e.g. Mareschal, 2007; Vandergrift, 2007; Brown et al., 2008; Chang, 2009; Webb and Chang, 2012), but the aim of teaching activities involving listening to songs is not always to comprehend. As pointed out by Farina (2013) and Jones (1999) depending on the language, having the orthographic form of the spoken word together with auditory input could hinder learning in terms of pronunciation. In fact, Mairano, Bassetti, Sokolovic-Perovic, and Cerni (2018, p.2) postulate that orthography could make L2 speakers produce a non-existing phonological category in the target language, confirming Murphey's (1990) hypothesis about foreign accent in adults:

> Reading done too soon as the main source of input, and hearing one's own voice subvocally, may partially account for the fact that many adults keep a strong non-native accent in a foreign language, while children, who are principally listeners not readers, seem to have little problem becoming native-like. (Murphey 1990, p.58)

Besides, bi-modal or multimodal input does not necessarily mean audio and orthography, it could refer to audio and visual sources together, therefore, a song's vocabulary can be learnt by providing extra support while listening, like pictures or actions (Medina, 1993). This idea was confirmed by de Vos, Schriefers, Mivard and Lemhöfer (2018) in their meta-analysis of incidental vocabulary learning from listening, where significant findings were always supported by L2 input with supporting materials (e.g. audiovisual materials such as videos).

2.5 *Another brick in the Wall (Pink Floyd, 1979)*:

Degree of directness in instruction

Ellis (2015) specified that SLA in instructed settings involves both a direct and indirect source of instruction. The former is defined as "providing learners with explicit information about the target of the instruction, often together with opportunities to practice the target" (p.241), while the latter as "setting up opportunities for learners to learn without specifying what the target of the instruction is" (p. 241). The literature sometimes refers to these two notions as explicit and implicit instruction (e.g. Bongaerts, van Summerin, Planken & Schils, 1997; Fullana, 2006; Kissling, 2008), and it should be clarified that for the present study the terms "explicit" and "implicit" will refer to the different types of instruction and should thus be considered as synonyms for "direct" and "indirect", without reference to explicit/implicit learning. Moreover, the learning fostered by these two types of instruction is going to be referred to as intentional in the case of direct instruction and incidental in the case of indirect instruction, following Hulstijn (2013). The author defines the former as a "deliberate attempt to commit factual information to memory" (Hulstijn, 2013, p.2632), while incidental learning is "the acquisition of a word or expression without the conscious intention to commit the element to memory" (Hulstijn, 2013, p.2632).

In their meta-analysis of L2 instruction in general, Norris and Ortega (2000) state that explicit instruction leads to immediate gains, but their sustainability is still debatable. In fact Hulstijn (2001, p.17), states that this short-term retention is possibly what makes researchers lean towards incidental vocabulary learning with more indirect means of instruction, such as extensive reading or listening, as "when L2 educationalists advocate incidental vocabulary learning while devaluating the role of intentional learning, what they probably mean is that the former procedure leads to information processing of a higher quality, and hence to better retention, than the latter procedure." A different view is shared by Ellis (2015), who maintains that the degree of retention depends on the linguistic feature, not on the type of instruction. In particular, he asserts that explicit instruction also has long-lasting effects on some linguistic features, while the implicit one benefits

discourse capabilities. Nevertheless, the necessarily limited nature of the direct mode as far as time is concerned, also leads to promoting a more implicit instruction.

Explicit pronunciation instruction is reported to be more helpful for adult L2 learners to achieve native-like pronunciation (Bongaerts et al., 1997; Fullana, 2006), but there are also studies that question its efficiency (Chung, 2008; Tominaga, 2009). In particular, Chung (2008) claims that "exposure and attention to the target feature were more relevant than other instructional characteristics for improving L2 speech". Such idea is supported by Kissling's (2013) study of L1 Spanish University learners of L2 English, where she compared explicit L2 phonetics instruction to a more implicit method (focused listening with dictation), and found that both types of instruction led to improvement, without significant differences between the two.

A similar situation occurs in vocabulary acquisition, as studied by Laufer (2006). The author demonstrated that in EFL contexts, vocabulary learning is assisted by explicit instruction. The author also added that in such settings meaning-focused instruction should also be integrated with an explicit form-oriented focus. There is evidence that vocabulary acquisition also takes place when an activity is not designed with a specific target (Gass, 1999; Hulstijn, 2001; Loewen, 2015) or incidentally with extensive reading, although at a slow rate (Pellicer-Sánchez & Schmitt, 2010; Serrano & Huang 2018; Waring & Takaki, 2003). On the contrary, pure extensive listening (EL) led to contrasting findings and is considered too challenging in an EFL context (e.g., Field, 2008) and that is why research about it is "nearly non-existent" (Chang, 2012, p. 26). In this study, Chang (2012) compared vocabulary learning and listening competence of EL and Intensive Listening (IL) in the RWL mode. The author concluded that while IL showed better gains in terms of vocabulary, EL led to significant better results in terms of listening skills, "implying that students' listening competence can be enhanced through reading and listening […] without formal instruction" (Chang 2012, p.39). Further studies by Webb and Chang confirmed such findings (Webb, S. & Chang, A, C-S., 2015a; Webb, S. & Chang, A, C-S., 2015b).

In relation to songs specifically, both Medina (1993) and Pavia et al. (2019) demonstrated that, although at small rates, incidental learning of vocabulary is

specifically possible by listening to songs with an indirect method of instruction, not only because of the slower speed of songs (Murphey, 1992), but also for the high degree of repetition of their vocabulary items (Pavia et al., 2019). Such positive findings in incidental vocabulary acquisition with songs were not confirmed by Maneshi (2017), who only observed significant results in spoken form and collocation recognition.

3. Aims and Research Questions

Considering the lack of experimental studies on the topic, the following study aims at exploring if and how listening to songs can provide a useful source for teaching pronunciation and vocabulary at the same time. Moreover, learning outcomes are going to be scrutinized in terms of perception/reception and production under different perspectives, both in terms of input modality (with or without lyrics) and type of instruction (direct or indirect). The two opposite modalities are going to be referred to as lyrics – no lyrics, since this is what distinguished the different groups who underwent treatment. In fact, all groups were exposed to audio and visual input, but one group did not have access to the lyrics, therefore such participants did not have access to the orthography of the target words.

The following research questions were addressed in this study:

1) Can the use of songs promote pronunciation and vocabulary learning in an EFL class?

2) Does input modality (with or without lyrics) have an effect on vocabulary and pronunciation learning through songs over time?

3) Does degree of directness in instruction (direct or indirect) have an effect on vocabulary and pronunciation learning through songs over time?

4. Methodology

4.1 Participants

The study was carried out in the Italian province of Monza-Brianza and it involved students born in 2006 attending the second year of middle school (*Scuola Media*). The school is a private religious comprehensive institution, ranging from pre-school to secondary school located in a 7000-inhabitants town, which is attended by mostly monolingual Italian students with high socio-economic status living in the surrounding area. At the time of data collection, the school had 700 students enrolled, 68 of which were in Year 8, the school year analysed. All the students from Year 8 were originally considered for the present study (N = 68); however, the final sample included 60 students. The eight students who did not take part in the study were either absent on the day of the treatment or had severe special education needs and their parents did not give consent for them to be involved. The parents of all participants signed a written consent to allow their children to participate in the study. In Year 8, there are three different classes and in order to ensure randomization of sampling, such intact classes were used as the three groups to undergo the treatment through different conditions. In terms of gender, overall there were 35 boys and 25 girls (group 1 with 20 students: 12 boys and 8 girls; group 2 with 18 students: 14 boys and 4 girls; group 3 with 22 students: 9 boys and 13 girls).

Students currently attend three hours of English as a foreign language (EFL) per week and one extra hour every two weeks to work on conversation or on improving their grammar knowledge. All lessons are taught by two native Italian teachers.

4.2 Design and Procedure

The design of the study is counterbalanced and includes a pre-test, treatment and an immediate post-test. Moreover, as far as perceptual/receptive skills are concerned, it also includes a delayed post-test one month after treatment.

The pre-test and the immediate post-test consist of two parts, with two sections each:

- Part I: Pronunciation Production and Vocabulary Production.

- Part II: Pronunciation Perception and Vocabulary Reception.

In the delayed post-test the students were only presented with Part II tests, due to time constraints. The order of completion was decided in order to avoid possible priming effects that could emerge when taking the perception/reception test before the production one. Participants took all the tests in the computer laboratory of the school, via the online form builder and creator JotForm (www.jotform.com; Ajmi, 2016). Nevertheless, due to issues with the school computers, the pronunciation production pre-test and post-test were recorded on an I-phone 8. In order to account for multimodality also during test times, a native English speaker was present at the immediate and delayed post-test, reading aloud the tests' questions.

Furthermore, after completing the post-tests, all participants completed a questionnaire in Italian with questions that tapped on the students' attitude towards the treatment. The questionnaire was devised in Italian to allow the students to be fully confident and eloquent in expressing their opinions. Unfortunately, because of space limitations it will not be possible to report on its results.

In order to assess the students' proficiency (the first part of the Quick Oxford Placement Test (OPT) and the 1,000, 2,000 and 3,000 levels of the monolingual version of the Vocabulary Levels Test (VLT, de Souza & Soares-Silva, 2015), available on Paul Nation's Website (http://www.victoria.ac.nz/lals/about/staff/paul-nation) were completed before the treatment. Thus, participants' proficiency was checked, in order to ensure that the three groups were comparable in that respect. The descriptive statistics for both tests, with mean and standard deviation in parentheses are reported in Table 1.

Table 1 – Descriptive statistics for proficiency tests

Group	Quick OPT /40	VLT /30
1	16.55 (4.180)	13.91 (5.088)
2	17.81 (3.371)	13.88 (5.691)
3	18.35 (2.914)	16.88 (3.951)

Two one-way ANOVAs were performed to compare the proficiency levels of the three groups both for the Quick OPT and the VLT. Results showed that there were no significant differences in terms of proficiency among groups for the Quick OPT, $F(2,41)=.942$, $p=.398$ or for the VLT, $F(2,41)=1.927$, $p=.159$, despite group 3's scores always being slightly higher.

Both the questionnaire, the Quick OPT and the VLT were taken in class in their paper versions, although the first one was collected during post-test and the other two during pre-test.

4.3 Treatment

Students underwent two days of treatment and each lesson lasted 50 minutes. As previously mentioned, the three groups were randomly assigned to three different conditions, in order to analyse the impact of different treatments on learning English through songs.

The songs chosen were two American pop singles from 2005: *Move Along* by The All-American Rejects and *Have a Nice Day* by Bon Jovi. Such songs were chosen because the students would not be familiar with them, had a catchy rhythm, and were likely to be considered appealing by the students. Besides, teachers were asked beforehand whether they had previously used such songs for classroom activities and they confirmed they had not.

Table 2 summarizes treatment for all three groups in terms of instruction and modality:

Table 2 - Summary of treatment

Group	Move Along	Have a Nice Day	Modality
1	direct pronunciation	direct vocabulary	lyrics
2	direct vocabulary	direct pronunciation	lyrics
3	direct pronunciation	direct vocabulary	no lyrics

The first day of treatment was dedicated to the song *Move Along* and Group 1 and 3 followed the lesson plan with a direct instructional focus on pronunciation. On the other hand, group two had a different lesson plan, since the focus was on direct vocabulary instruction. The detailed lesson plans are available in Appendix A.

14

The second day was dedicated to *Have a Nice Day* and the lessons were performed following the same plans reported above, only the other way around. Thus, Group 2 followed the lesson plan with an explicit instruction focus on pronunciation, while Groups 1 and 3 followed the explicit vocabulary one. During both days, regardless of instruction focus or modality, all participants listened to the song four times during treatment.

Despite having a dissimilar focus and different activities, both lesson plans followed the same global structure, as the students first watched the official video-clips of the songs without on screen or paper lyrics. Then, after ice-breaker activities, Groups 1 and 2 were given the lyrics to the song, where the target words directly taught were highlighted in red (Appendix B). In the case of Group 3 (no-lyrics modality), they were handed in a worksheet with images taken from Google rather than with the words of the song, where the target items stood out because of a green band next to the image (Appendix C). All groups then listened to the song following the allocated worksheet. Then, after some activities on the target forms, which included one further listening with access to the lyrics/no-lyrics worksheets, there was a plenary with a sing-along activity.

4.4 Target Words

In terms of pronunciation, both songs presented 7 different items of two specific pronunciation features: aspirated /h/ and English dental fricatives. Such sounds were chosen as topic of analysis for pronunciation, since they are not present in Italian and might therefore be problematic for Italian speakers Wheelock (2016). As for the former feature, Italian speakers tend to fall into an h-deletion pattern when pronouncing English words, while for the latter they have the tendency to form the labial-dental fricative [f] for /θ/ and the laminal denti-alveolar [d] for /ð/. Table 3 presents the target words for pronunciation with their frequency within the song and according to the British National Corpus/ Corpus of Contemporary American English (BNC/COCA) word family lists. One of the target words, *ahead*, is reported with an asterisk, since it does not present an aspirated /h/ at the beginning of the word. Despite such difference, it nevertheless presents the same possible h-deletion feature and was thus included as a target item.

Table 3 - Target Words for Pronunciation

	Target Words Pronunciation	Frequency Within Song	Frequency Level
Move Along	ahead*	1	K-1
	had	1	K-1
	held	2	K-1
	hands	4	K-1
	hold	2	K-1
	hope	5	K-1
	heart	1	K-1
Have a Nice Day	the	23	K-1
	there	3	K-1
	thing	3	K-1
	through	3	K-1
	that	4	K-1
	with	3	K-1
	nothing	1	K-1

As far as vocabulary is concerned, the following target words were selected:

Table 4 - Target Words for Vocabulary

	Target Words Vocabulary	Frequency Within Song	Frequency Level
Move Along	to waste	1	K-1
	fill	1	K-1
	to fall	1	K-1
	sinking	1	K-2
	to move along	29	K-1
	to make it through	5	K-1
	deceiving	1	K-6
	to right	4	K-1
Have a Nice Day	shining	3	K-2
	standing	3	K-1
	dice	3	K-5
	ledge	3	K-6
	to stand my ground	1	K-1
	brave	1	K-2
	to take a stand	1	K-1
	dead-end street	1	K-1

Because one of the groups did not have access to the lyrics, the main criteria in the choice of target words was that items had to be easily identifiable with an image and that the students were not expected to know, based on their syllabus.

The lyrics of each song were analysed with Vocabulary Profilers (www.lextutor.ca), which breaks texts down by word frequencies according to

the British National Corpus/ Corpus of Contemporary American English (BNC/COCA) word family lists:

Table 5 - Move Along Lexical Profile

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. token (%)
K-1 Words :	64 (97.0)	75 (96.15)	323 (99.4)	99.4
Coverage 95 ?				
K-2 Words :	1 (1.5)	1 (1.28)	1 (0.3)	99.7
Coverage 98				
K-3 Words :				
K-4 Words :				
K-5 Words :				
K-6 Words :	1 (1.5)	1 (1.28)	1 (0.3)	100.0

Table 6 - Have a Nice Day Lexical Profile

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. token (%)
K-1 Words :	96 (91.4)	111 (90.98)	343 (95.0)	95.0
Coverage 95 ?				
K-2 Words :	5 (4.8)	5 (4.10)	7 (1.9)	96.9
K-3 Words :				
K-4 Words :	1 (1.0)	1 (0.82)	3 (0.8)	97.7
K-5 Words :	2 (1.9)	2 (1.64)	4 (1.1)	98.8
Coverage 98				
K-6 Words :	1 (1.0)	1 (0.82)	3 (0.8)	99.6

Therefore, in the case of *Move Along*, in order to understand 95% of words in the song, it was enough to know the first 1000 most frequent word families and to reach the 98% threshold of understanding the first 2000 most frequent word families. Likewise, for *Have a Nice Day*, to understand 95% of words in the song, it was enough to know the first 1000 most frequent word families, while to get to 98% coverage it was necessary to know the first 5000.

4.5 Tests

All tests completed by the students are presented in Appendices D (Production) and E (Perception/Reception). As for pronunciation production (Appendix D), the students were recorded with an I-phone 8 saying all 14 pronunciation target words in a word list reading. Then, due to time constraint issues, only two items (i.e. "through" and "nothing") were considered for the analysis, since they offered the same feature, the voiceless dental fricative [θ],

in two different positions, one at the beginning and the other at the end of the word. Besides, among all the target words, "through" was the one the students were less likely to know and therefore would be a better litmus test for the effectiveness of the treatment. Such sound files were extracted from each participant's recording, their background noise was cancelled and their volume was normalised. Such stimuli were then presented to four non-trained native American-English-speaking raters through Praat, with a rating task based on a 7-point Likert scale (1=extremely inaccurate to 7=extremely accurate). The raters were two male (25 and 30 years old) and two female (28 and 32 years old) teachers of English, who did not have direct familiarity with the Italian accent. Both pre-test and post-test stimuli were presented in random order in the same task, so as to avoid raters knowing whether the audio file came from before or after the treatment. The results for each participant were obtained by calculating the mean among the scores given by the four raters and computing an average result for pre-test and another for post-test for each of the two target items. Each rater was presented with 240 stimuli to assess individually through Praat and it took them around 35 minutes to complete the task.

In terms of pronunciation perception, the participants had to perform an AX auditory discrimination task (Rallo Fabra, 2016) with 14 items (Appendix E), where they had to decide whether the two words they heard were the same or different. The 14 target words were recorded by an American English native speaker in laboratory conditions embedded in a sentence, and then they were extracted and their volume was normalized through Praat (Boersma, 2002). Then two stimuli were put together to form a continuum with a pause in the middle and uploaded on JotForm. After the students submitted the task, an automatic e-mail with all their answers was generated. In case their answer was correct, the reply was coded as 1, whereas if incorrect, it was coded as 0.

The vocabulary production test (Appendix D) was a C-test, which, as reported by Singleton (1999), "is a reduced-redundancy procedure requiring the testee, essentially, to restore to wholeness a text nearly 50 per cent of whose constituent words have had their second half deleted". Thus, the participants were presented with half of the word embedded in a sentence and they had to complete it. Overall, there were 16 items in the test and in case the

18

answer was reported correctly, it was coded as 1, whereas if incorrect or left blank, it was coded as 0.

The receptive vocabulary knowledge test (Appendix E) was a multiple-choice test with 16 items, where students had to choose the correct definition for each target word. The options were based on antonym forms or on words that had similar spelling but different meaning. Most definitions were taken from the MacMillan Dictionary (Rundell, 2002). Moreover, participants were presented with an "I don't know" option. If their answer was correct, the reply was coded as 1, whereas if incorrect or unknown, it was coded as 0.

4.6 Statistical Analyses

Since none of the accuracy measures had a normal distribution, four Generalized Linear Mixed Models (GLMMs) were run using SPSS v25 with a repeated measures Diagonal structure, a random intercept for Subject and each of the four measures as dependent variables (Pronunciation Production Accuracy, Pronunciation Perception Accuracy, Vocabulary Reception Accuracy and Vocabulary Production Accuracy). The first dependent variable was a scale modelled through an inverse Gaussian, while the remaining three had a binomial distribution modelled with a Logit link. All models included Subject and Item as random factors, and the same fixed effects structure, i.e., Time (pre-test, immediate post-test and delayed post-test for perception tests; pre-test and immediate post-test for production tests), Modality (Lyrics, No Lyrics), Instruction (Direct, Indirect), with their paired (Modality*Time; Instruction*Time; Modality*Instruction) and three-way (Modality*Instruction*Time) interactions. Sixteen participants were excluded from the statistical analysis as outliers or as children with special education needs (SEN).

5. Results

The results are presented in two different sections: Pronunciation and Vocabulary. Moreover, each section is divided into two subsections, to comment on the participants' outcome in terms of perceptual/receptive and productive skills. At the beginning of each results section, a table of the

descriptive statistics is reported with the mean together with the standard deviation in brackets. In this by-item data-file, the variables Pronunciation_Perception_Accuracy[2], Vocabulary_Production_Accuracy and Vocabulary_Reception_Accuracy are categorical ones, where the minimum value in the descriptive statistics is 0 and the maximum is 1, while Pronunciation_Production_Accuracy is a scale from 1 to 7.

5.1 Pronunciation

5.1.1 Productive Skills

The descriptive statistics (means and SD in parentheses) for the Pronunciation Production Accuracy test for the different conditions are shown in Table 7, in which it is possible to see that both in the pre-test and in the post-test, students who underwent the treatment without lyrics performed better than the ones who had access to them. Table 7 shows that after the treatment there were gains for the lyrics modality and the indirect instruction mode, while there were slight losses for the no-lyrics modality and the direct mode. The immediate gains were calculated subtracting the pre-test results from the post-test ones.

Table 7 - Descriptive Statistics Pronunciation Production Accuracy

	Pre-test /7	Post-test /7	Immediate Gains
Lyrics	3.21 (1.09)	3.46 (1.12)	.25
No lyrics	3.66 (0.91)	3.53 (1.10)	-.13
Direct	3.51 (0.96)	3.43 (1.11)	-.08
Indirect	3.17 (1.16)	3.60 (1.11)	.43

The results of the GLMM for Pronunciation Production Accuracy appear in Table 8, in which significant effects are reported with asterisks next to the value (* when $.0.5 < p > .01$, ** if p $< .01$ and *** if $p < .001$). A larger version of Tables 8, 10, 12 and 14 is also available in Appendix F. Figure 1 provides a visual representation of the results. The thicker the line tracing back to Accuracy, the stronger the effect of such independent variable on the dependent one. Thin lines identify non-significant fixed effects.

---

[2] Capital letters are used when referring to specific variables or tests in the study.

Fixed Effects

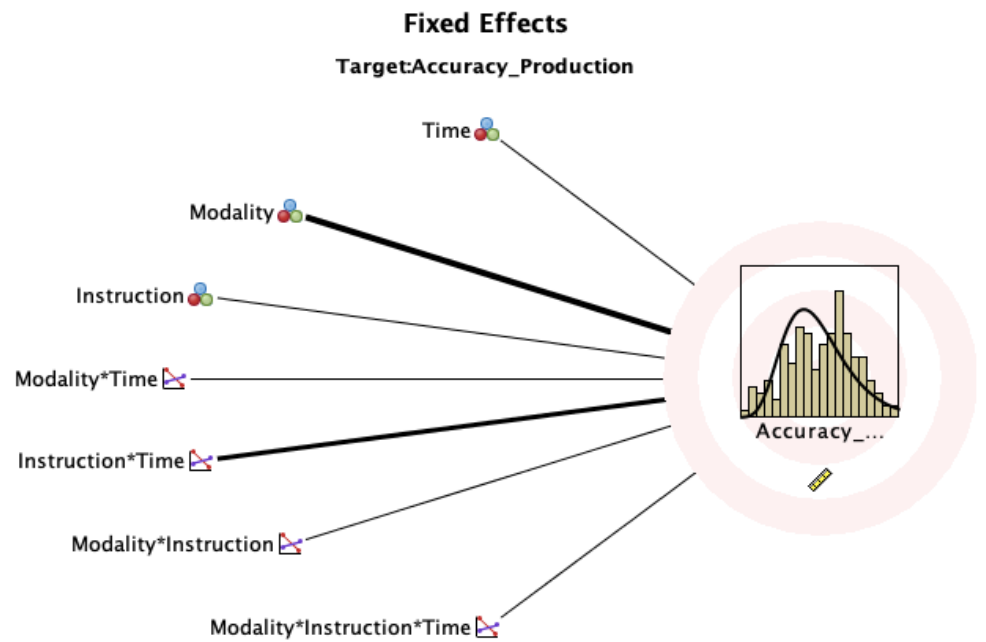**Target:Accuracy_Production**

Figure 1: Pronunciation Production Accuracy GLMM

The graph shows a non-significant effect of Time, a significant effect of Modality with the no-lyrics condition performing significantly better than the lyrics one and a non-significant effect of Instruction. In terms of interactions, there is none between Modality*Time or between Instruction*Time. Moreover, there is no interaction among Modality*Instruction*Time. Although it was not possible to compute Modality*Instruction, due to the fact that there was an empty set (direct instruction with no-lyrics modality), Table 8 shows that when examining the effect of Modality for each condition separately, it was found that with indirect instruction there was a significant difference between the no-lyrics and the lyrics condition, suggesting higher scores in connection with the no-lyrics modality.

Table 8 - Pronunciation Production results

Source	F	df1	df2	Sig.	Time				Modality				Degree of Directness		
Time	.106	1	166	.745	T2-T1	.012	.927								
Modality	6.504	1	166	.012*					No Lyrics – Lyrics	.346		.026*			
Instruction	.517	1	166	.473									Direct – Indirect	-.083	.536
Modality*Time	.147	1	166	.702	Lyrics	T2-T1	.141	.377							
					No Lyrics	T2-T1	-.274	.223							
Instruction *Time	3.074	1	166	.081	Indirect	T2-T1	-.198	.230							
					Direct	T2-T1	.421	.072							
Modality*Instruction	Unable to compute due to 0 degrees of freedom								No Lyrics –Lyrics	Indirect	.401	.028*			
									No Lyrics –Lyrics	Direct	n/a	n/a			
Modality* Instruction *Time	Unable to compute due to 0 degrees of freedom				Lyrics Indirect	T2-T1	-.130	.574							
					No Lyrics Indirect	T2-T1	-.274	.223							
					Lyrics Direct	T2-T1	.421	.072							
					No Lyrics Direct	T2-T1	n/a	.n/a							

5.1.2 Perceptual Skills

The descriptive statistics for Pronunciation Perception Accuracy are presented in Table 9, which shows that the students who underwent the treatment with lyrics improved between the pre-test and the immediate post-test, but that such gains did not last until the delayed post-test. On the other hand, the no-lyrics group demonstrated improvement both in the immediate and even more in the delayed post-test, showing higher long-term gains. As far as Instruction is concerned, the indirect mode follows the same path as the lyrics one, showing an increase at immediate post-test, which remarkably reduces by delayed post-test. The direct mode registers gains between pre-test and the immediate post-test, which then stay stable until delayed post-test. The immediate gains were calculated subtracting the pre-test results from the immediate post-test ones, while the long-term gains were calculated deducting pre-test scores from delayed post-test ones.

Table 9 - Descriptive Statistics Pronunciation Perception Accuracy

	Pre-test /1	Post-test /1	Delayed post-test /1	Immediate Gains	Long-Term Gains
Lyrics	.73 (.44)	.79 (.41)	.74 (.44)	.06	.01
No lyrics	.76 (.43)	.82 (.39)	.85 (.35)	.06	.09
Direct	.73 (.45)	.80 (.40)	.80 (.40)	.07	.07
Indirect	.76 (.43)	.80 (.40)	.77 (.42)	.04	.01

Figure 2 and Table 10 show the results for the Pronunciation Perception Accuracy GLMM test, indicating that Time and Modality have a significant fixed effect on the dependent variable, while Instruction does not.
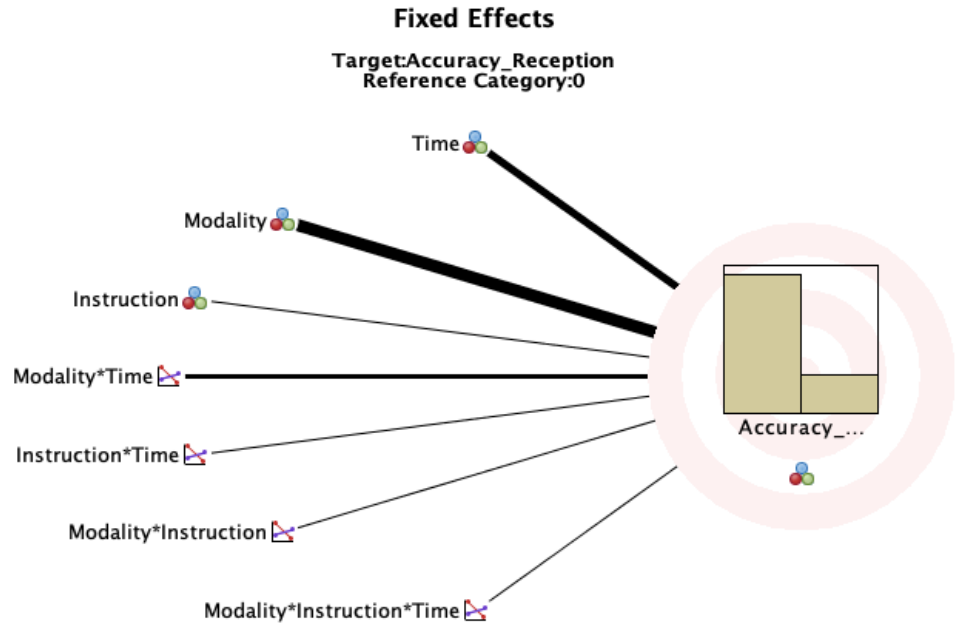

Figure 2 - Pronunciation Perception Accuracy GLMM

In terms of Time, scores were higher in the immediate post-test and delayed post-test, than in the pre-test, indicating that songs were helpful to promote perceptual pronunciation skills. In terms of Modality, the no-lyrics group performed better than the lyrics one.

There is not a significant Modality*Time interaction, indicating that the effect of the treatment with and without lyrics was equivalent. When examining the effect of Time for each condition separately, it was found that, in the lyrics condition, there were significant gains between pre and immediate post-test, whereas for the no-lyrics condition the significant difference was between pre and delayed post-test, suggesting longer term gains for this condition.

Likewise, there is no Instruction*Time interaction, suggesting that the effect of treatment with direct or indirect instruction was equal, although when analysing the effect of Time for the two conditions separately, there were significant gains in the immediate post-test, which were also retained at the

delayed post-test in the case of the direct condition, but not for the indirect one.

Similarly, Modality*Instruction do not interact, but when comparing the effect of Modality on the two separate types of instruction, there was a significant difference between the no-lyrics and the lyrics group in the indirect mode, with the no-lyrics group scoring significantly higher than the lyrics one. Despite the fact that there is no Modality*Instruction*Time effect, examining the effect of Time for each condition separately allows for significantly higher scores to be registered for the no-lyrics modality with direct instruction between delayed post-test and pre-test, arguing for higher long-term perceptual gains for students not having access to the lyrics and receiving explicit training.

Table 10 - Pronunciation Perception results

Source	F	df1	df2	Sig.	Time				Modality				Degree of Directness		
Time	4.804	2	1,794	.008**	T2-T1		.059	.020*							
					T3-T1		.059	.026*							
					T3-T2		-.000	.986							
Modality	12.313	1	1,794	.000***					No Lyrics - Lyrics		.063	.006**			
Instruction	.010	1	1,794	.921									Direct - Indirect	.002	.921
Modality*Time	2.389	2	1,794	.092	Lyrics	T2-T1	.067	.042*							
						T3-T1	.017	.603							
						T3-T2	-.049	.124							
					No Lyrics	T2-T1	.053	.119							
						T3-T1	.085	.021*							
						T3-T2	.032	.245							
Instruction*Time	.955	2	1,794	.385	Indirect	T2-T1	.041	.210							
						T3-T1	.027	.421							
						T3-T2	-.013	.673							
					Direct	T2-T1	.079	.025*							
						T3-T1	.091	.016*							
						T3-T2	.011	.681							
Modality* Instruction	.740	1	1,794	.390					No Lyrics –Lyrics	Indirect	.082	.014*			
									No Lyrics –Lyrics	Direct	.044	.115			
Modality* Instruction *Time	.606	3	1,794	.611	Lyrics Indirect	T2-T1	.070	.118							
						T3-T1	-.000	1.000							
						T3-T2	-.070	.135							
					No Lyrics Indirect	T2-T1	.017	.686							
						T3-T1	.042	.335							
						T3-T2	.024	.546							
					Lyrics Direct	T2-T1	.064	.145							
						T3-T1	.033	.473							
						T3-T2	-.031	.459							
					No Lyrics Direct	T2-T1	.095	.060							
						T3-T1	.134	.013*							
						T3-T2	.040	.278							

5.2 Vocabulary

5.2.1 Productive Skills

The descriptive statistics for Vocabulary Production Accuracy are laid out in Table 11 and they show gains between pre-test and immediate post-test for all conditions.

Table 11 - Descriptive Statistics Vocabulary Production Accuracy

	Pre-test /1	Post-test /1	Immediate Gains
Lyrics	.07 (.25)	.42 (.49)	.35
No lyrics	.09 (.28)	.51 (.50)	.42
Direct	.07 (.25)	.56 (.50)	.49
Indirect	.09 (.28)	.36 (.48)	.27

Figure 3 presents the Vocabulary Production Accuracy GLMM results (all values reported in Table 12). The graph shows a significant effect of the independent variable Time, with significant gains between pre and post-test, but no effect of Modality or Instruction.
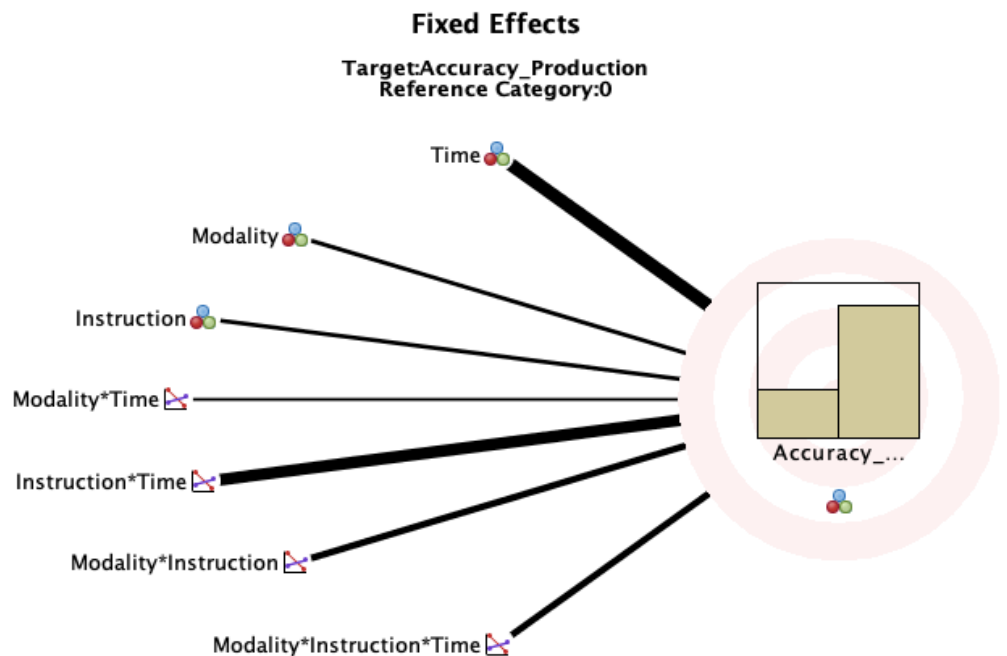

Figure 3 - Vocabulary Production Accuracy GLMM

Modality*Time do not interact, which means that the two modalities of treatment do not lead to significantly different performances among

participants. However, when the effect of Modality at different times is analysed, it is shown that the no-lyrics condition scores were significantly higher than the lyrics ones in the post-test, while no differences existed in the pre-test.

Instruction*Time interact and in order to further explore such interaction, a pairwise comparison was carried out and it turned out that Time had an effect, both in the case of the indirect and the direct mode. Whereas the same pairwise comparison, conducted to study the effect of Instruction at different times, showed that the direct instruction led to significantly better results in the post-test than the indirect one.

Modality*Instruction also interact, and the pairwise comparison reveals that with indirect instruction, the no-lyrics group reached significantly higher scores.

Modality*Instruction*Time also interact. In terms of pairwise comparisons, the effect of Time is significant on all combinations of conditions (Lyrics*Indirect, No Lyrics*Indirect, Lyrics*Direct, No Lyrics*Direct), with significantly higher scores in the post-test. When the focus is on the effect of Instruction, higher scores are registered for both the no-lyrics and the lyrics conditions at post-test time in relation to direct instruction. Besides, with indirect instruction, participants of the no-lyrics group performed significantly better in the pre-test.

Table 12 - Vocabulary Production results

Source	F	df1	df2	Sig.	Time				Modality				Degree of Directness				
Time	239.901	1	1,400	.000***	T2-T1	.391	.000***										
Modality	2.507	1	1,400	.114					No Lyrics - Lyrics	.0039	.129						
Instruction	2.338	1	1,400	.126									Direct - Indirect	.041	.139		
Modality* Time	0.409	1	1,400	.523	Lyrics	T2-T1	.351	.000***	T1	No Lyrics – Lyrics	.009	.590					
					No Lyrics	T2-T1	.432	.000***	T2	No Lyrics – Lyrics	.089	.033*					
Instruction* Time	14.731	1	1,400	.000***	Indirect	T2-T1	.273	.000***					T1	Direct – Indirect	-.020	.223	
					Direct	T2-T1	.512	.000***					T2	Direct – Indirect	.219	.000***	
Modality* Instruction	4.220	1	1,400	.040*					No Lyrics – Lyrics	Indirect	.090	.014*					
									No Lyrics – Lyrics	Direct	-.022	.608					
Modality* Instruction* Time	3.606	2	1,400	.017*	Lyrics Indirect	T2-T1	.259	.000***	T1 Indirect	No Lyrics – Lyrics	.063	.030*	T1 Lyrics	Direct – Indirect	.023	.227	
					No Lyrics Indirect	T2-T1	.277	.000***	T2 Indirect	No Lyrics – Lyrics	.081	.181	T2 Lyrics	Direct – Indirect	.213	.000***	
					Lyrics Direct	T2-T1	.449	.000***	T1 Direct	No Lyrics – Lyrics	-.031	.139	T1 No Lyrics	Direct – Indirect	-.071	.024*	
					No Lyrics Direct	T2-T1	.569	.000***	T2 Explicti	No Lyrics – Lyrics	.089	.147	T2 No Lyrics	Direct – Indirect	.221	.003**	

5.2.2 Receptive Skills

Table 13 shows the descriptive statistics results for the Vocabulary Reception test, which shows a similar trend in all modalities. Participants show improvement between pre-test and immediate post-test and then a subsequent decrease at delayed post-test time. However, there are still long-term gains in all conditions.

Table 13 - Descriptive Statistics Vocabulary Reception Accuracy

	Pre-test /1	Post-test /1	Delayed post-test /1	Immediate Gains	Long-term Gains
Lyrics	.25 (.43)	.43 (.49)	.36 (.48)	.18	.11
No lyrics	.25 (.43)	.49 (.50)	.46 (.50)	.24	.21
Direct	.26 (.44)	.53 (.50)	.46 (.50)	.27	.20
Indirect	.23 (.42)	.37 (.48)	.34 (.47)	.14	.11

Figure 4 characterises the GLMM Vocabulary Reception Accuracy results (all values available in Table 14). Time, Modality and Instruction all have a significant fixed effect on the accuracy results, as it is possible to see from the thickness of the lines connecting such variables to Accuracy in the figure below.
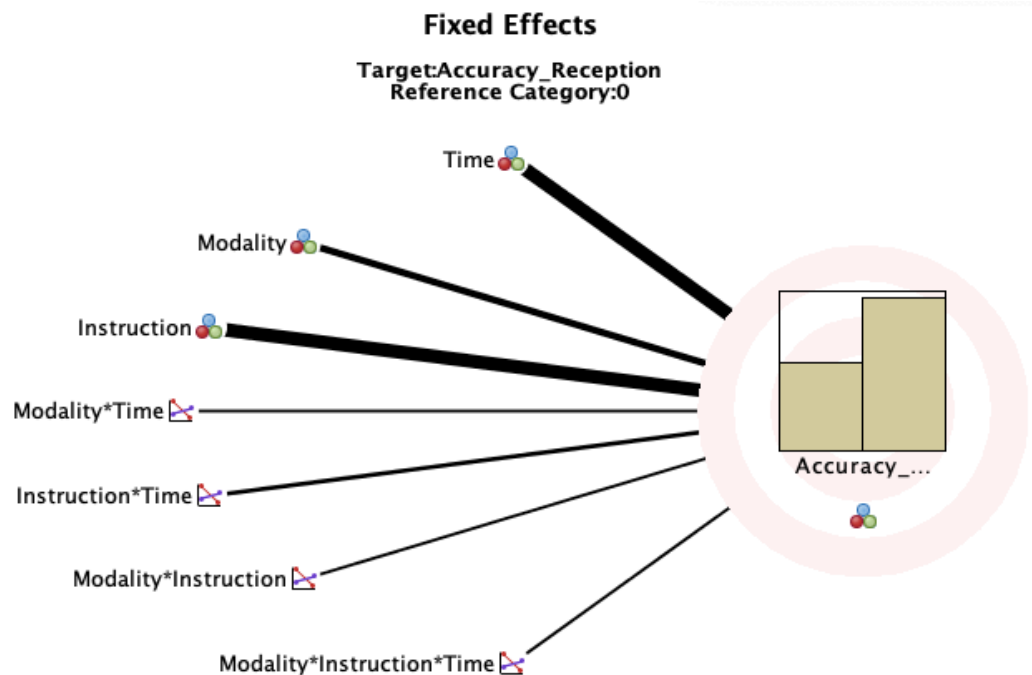

Figure 4 - Vocabulary Reception Accuracy GLMM

As for Time, immediate and delayed post-tests register higher scores than pre-test, whereas the no lyrics modality is connected to better results for Modality and the direct mode for Instruction.

There is no significant Modality*Time interaction, indicating that there are no differences in gains in receptive vocabulary knowledge between the lyrics and the no-lyrics modalities. When examining the effect of Time for each condition separately, it was found that, in the lyrics condition, there were significant gains between pre and both the immediate and the delayed post-test, but there were also significant losses between immediate and delayed post-test. On the other hand, such decrease is non-significant for the no-lyrics group, suggesting that although there are long-term gains for both conditions, the no-lyrics one tends to preserve more content through time. In fact, if the focus of the pairwise comparison is on the effect of Modality, in the delayed post-test the no-lyrics group performed significantly better than the lyrics one.

Moreover, there is no Instruction*Time interaction, so the treatment was equally effective with direct or indirect instruction. When analysing the effect of Time for the two separate conditions, there were significant short and long-term gains for both types of instructions. There were also non-significant losses between immediate and delayed post-test times for the two conditions. If the focus of the analysis is on the effect of Instruction, though, the direct mode is shown to lead to better short and long-term results.

Modality*Instruction do not significantly interact between themselves or in their three-way comparison with Modality*Instruction*Time. However, when analysing the effect of Time in the latter three-way pairwise comparison, the no-lyrics indirect, lyrics direct and no-lyrics direct conditions lead to significantly better results at both immediate and delayed post-test times, while the lyrics indirect combination only does so in the immediate post-test. When the focus of the pairwise comparisons is switched to the effect of Instruction, the lyrics direct condition was linked to significantly better scores both in the immediate and the delayed post-test, while the no-lyrics direct one is only linked to higher long-term gains.

Table 14 - Vocabulary Reception results

Source	F	df1	df2	Sig.	Time				Modality				Degree of Directness			
Time	34.837	2	2,100	.000***	T2-T1	.214		.000***								
					T3-T1	.165		.000***								
					T3-T2	-.049		.078								
Modality	5.744	1	2,100	.017*					No Lyrics - Lyrics		.053	.018*				
Instruction	16.803	1	2,100	.000***									Direct - Indirect		.101	.000***
Modality* Time	1.424	2	2,100	.241	Lyrics	T2-T1	.186	.000***	T1	No Lyrics – Lyrics	.004	.903				
						T3-T1	.116	.000***	T2	No Lyrics – Lyrics	.060	.129				
						T3-T2	-.070	.041	T3	No Lyrics – Lyrics	.103	.008**				
					No Lyrics	T2-T1	.242	.000***								
						T3-T1	.215	.000***								
						T3-T2	-.027	.542								
Instruction* Time	2.012	2	2,100	.134	Indirect	T2-T1	.151	.000***					T1	Direct – Indirect	.031	.360
						T3-T1	.122	.001**								
						T3-T2	-.029	.445					T2	Direct – Indirect	.159	.000***
					Direct	T2-T1	.279	.000***								
						T3-T1	.210	.000***					T3	Direct – Indirect	.119	.003**
						T3-T2	-.069	.080								
Modality* Instruction	.009	1	2,100	.908					No Lyrics –Lyrics	Indirect	.055	.074				
									No Lyrics –Lyrics	Direct	.050	.120				
Modality* Instruction* Time	.112	3	2,100	.857	Lyrics Indirect	T2-T1	.113	.010*					T1 Lyrics	Direct – Indirect	.019	.642
						T3-T1	.065	.122								
						T3-T2	-.047	.296					T2 Lyrics	Direct – Indirect	.170	.000**
					No Lyrics Indirect	T2-T1	.190	.001**								
						T3-T1	.182	.001**					T3 Lyrics	Direct – Indirect	.125	.008**
						T3-T2	-.008	.899								
					Lyrics Direct	T2-T1	.264	.000***					T1 No Lyrics	Direct – Indirect	.043	.436
						T3-T1	.171	.000***								
						T3-T2	-.092	.060					T2 No Lyrics	Direct – Indirect	.148	.027*
					No Lyrics Direct	T2-T1	.294	.000***								
						T3-T1	.249	.000***					T3 No Lyrics	Direct – Indirect	.110	.103
						T3-T2	-.046	.457								

Thus, the summary of the significant fixed effects and interactions is reported in Table 15:

Table 15 - Summary of Significant Fixed Effects and Interactions

	Time	Modality	Instruction	Modality*Time	Instruction*Time	Modality*Instruction	Modality* Instruction*Time
Pronunciation Production		* (No Lyrics)					
Pronunciation Perception	** (T2 and T3)	*** (No Lyrics)					
Vocabulary Production	*** (T2)				*** (Indirect T2 – Direct T2)	* (No Lyrics Indirect)	** (Lyrics Direct T2; No-Lyrics Direct T2; Lyrics Indirect T2; No-Lyrics Indirect T2)
Vocabulary Reception	*** (T2 and T3)	* (No Lyrics)	*** (Direct)				

6. Discussion

The analyses carried out now allow for the abovementioned research questions to be answered:

1) **Can the use of songs promote pronunciation and vocabulary learning in an EFL class?**

The Pronunciation Perception, Vocabulary Production and Vocabulary Reception results showed that indeed pronunciation and vocabulary can be learned through songs. The Perception/Reception tests showed that there were gains both at immediate post-test and delayed post-test time, which leads to believe that such gains are kept long-term. These results confirm Farina (1993) and endorse the validity of classroom activities that go beyond comprehension to develop other L2 skills (Arnold & Herrick, 2017).

The only area where treatment did not seem to be successful was Pronunciation Production. This outcome is in line with research from Christiner and Reiterer (2015), who highlighted that singing is beneficial to attain a more accurate L2 pronunciation, but the present treatment did not entail singing or vocal motor training. Nevertheless, it should also be noted that the short duration of treatment might have limited improvement only to perception, but a longer period of exposure could have led to a production one, following what was postulated by the SLM (Flege, 1995).

Furthermore, the fact that the no-lyrics group showed improvement between immediate and delayed post-test in the Pronunciation Perception test suggests that possibly these students had higher aptitude at learning through songs and it should also be noted though that, despite not statistically significant, their performance in the proficiency tests was better than the one of the other groups. It would also be interesting to see if not having access to the words enhanced the *song-stuck-in-my-head* phenomenon (Murphey, 1992).

2) **Does input modality (with or without lyrics) have an effect on pronunciation and vocabulary learning through songs over time?**

Modality did not stand out as a significant variable over time in any of the tests presented, as there were no Modality*Time interactions. This finding confirms that learning can occur both with and without access to orthography

for pronunciation (e.g. Farina, 1993) and for vocabulary (e.g. Medina, 1993).

The pairwise comparisons also allow for deeper analyses to be carried out. While the Pronunciation Production Modality*Time interactions showed no significant result, the Pronunciation Perception pairwise comparisons show that whereas the lyrics modality groups performed significantly better at immediate post-test time, the no-lyrics one did so at delayed post-test time. The pairwise comparisons of the Vocabulary Reception test paint a similar picture, since although the test presented significant interactions for the two modalities at both post-test times, the lyrics modality also registered a significant loss between immediate and delayed post-test. These findings support the idea that introducing new content through another semiotic system, such as music or images, helps students make deeper associations (Olshansky 2018; Albers, 2007; Leland & Harste, 1994).

As far as Vocabulary Production is concerned, the pairwise comparisons show that when analysing the effect of Time for the two separate conditions, there were significant gains for both types of modality. When the effect of Modality is under scrutiny, though, the no-lyrics modality shows significantly better results than the lyrics one. Such result hypothesises the potential applicability of Medina's findings in terms of receptive skills to productive ones as well. In terms of vocabulary acquisition, the literature suggests that the lyrics modality should have led to higher gains, due to the RWL circumstance (e.g. Chang, 2009). Nevertheless, the abovementioned studies did not involve RWL to a song, but rather guided reading of books aimed at the development of both comprehension and listening skills, while Medina's (1993) study focused on vocabulary acquisition through songs, although with younger learners. Based on Medina's and the present study's results, songs should be analysed as a separate RWL subgroup, especially when the focus of the teaching activity is not comprehension.

3) Does instruction (direct or indirect) have an effect on pronunciation and vocabulary learning through songs over time?

Instruction was a significant variable over time only in the Vocabulary Production test. When analysing the effect of instruction on the two time frames, the direct mode of instruction leads to a significantly better

performance by students, while when the focus is on the effect of time, both instruction modes are significant. Although the Vocabulary Reception test does not show a significant Instruction*Time interaction, the pairwise comparisons follow the same pattern, with the direct instruction leading to better results when analysing the effect of instruction. These findings are in line with literature supporting the effectiveness of direct instruction in vocabulary acquisition (Beck, McKeown, & Kucan, 2002; Laufer, 2006; Hindman & Wasik, 2013; Norris & Ortega, 2000).

As far as Pronunciation Production is concerned, instruction had no effect on students' results, which confirms Chung's (2008) findings that there are more salient features that influence improvement in L2 speech. On the other hand, as for the Pronunciation Perception results, there were higher scores connected to the direct mode of instruction. In fact, although Instruction*Time did not reach significance (same as in the Production test), the pairwise comparison in the perception test shows higher scores connected to the direct instruction, both at immediate and at delayed post-test time (e.g. Bongaerts et al., 1997; Fullana, 2006).

The fact that the type of instruction was not critical for learning to occur in the Perception/Reception tests can be considered a supporting argument not only for incidental receptive learning of vocabulary through songs, as in Pavia et al. (2019), but also for the effectiveness of extensive listening through songs, if reinforced with images, (de Vos et al., 2018).

In addition to the findings in relation to the research questions, from the Modality*Instruction and Modality*Instruction*Time, there was a significant interaction between the no-lyrics modality and the indirect type of instruction in all tests. Such result supports not only a less verbocentric outlook on the educational system (Olshansky, 2018), but also the effectiveness of extensive listening, when the input is targeted not beyond the *i*+1 comprehensible input threshold (Krashen, 1982) and through a song supported with images.

7. Limitations and Further Research

The first limitation of the present study is that it was carried out using just two songs and treatment only lasted two days. Besides, Pellicer-Sánchez (2018) says that one-shot design studies such as this one tend to give immediate gains that do not last, therefore longitudinal studies would be beneficial to solve these issues. Moreover, due to the number of variables involved, the statistical analyses would have benefitted from a higher number of participants, in order for results to be generalizable. SEN children's data was collected but not analysed at the moment, in order to reach broader conclusions. Nevertheless, the analysis of their performance will be valuable to check if children with dyslexia performed better in the no-lyrics modality. Further research could also compare attainment through a lesson with songs and a more traditional one, both in terms of Pronunciation and Vocabulary.

Furthermore, due to time limitations, the frequency of words within songs was not taken into consideration. Such variable turned out to be crucial in Pavia et al. (2019) and thus deserves further attention. Likewise, the qualitative analysis involving students' and teachers' perceptions could not be included at the present time. Besides, due to technical difficulties, the Pronunciation Production test had to be recorded on an I-phone, thus lacking quality, and was not analysed in its entirety.

8. Conclusions and Pedagogical Implications

This study shows that it is possible to learn both L2 pronunciation and vocabulary by listening to L2 songs. Although listening activities can be challenging for students, their success also depends on the target of the activity, which does not necessarily need to be comprehension (Chang, 2009; Arnold & Herrick, 2017). In particular, since pronunciation is an area teachers do not feel confident approaching (Darcy, 2018), songs can provide a supportive authentic material for instructors, especially since in the EFL context there is an elevated number of non-native-English teachers (Copland et al., 2014) who might not feel confident in teaching L2 pronunciation. In particular, since in Italy Music is a compulsory subject up until Year 9, this

would be a precious opportunity to carry out a Content and Language Integrated Learning (CLIL) project, thus relying on the expertise of Music teachers for the vocal motor training necessary to improve Pronunciation Production skills (Christiner & Reiterer, 2015). Although the subject of the 2015 study were singers and musicians, Richards and Durrant (2003) demonstrated that people only need the opportunity to develop their singing skills and that with the right combination of factors, anybody can learn how to sing. Furthermore, in this way the intended aim of CLIL, with 50% focus on content (vocal training) and 50% on language (L2 vocabulary) would also be made a reality, without having too high expectations for subject teachers to carry out a lesson in a foreign language or for language teachers to become experts in other subjects.

This study also suggests that vocal training would be as effective if carried out without lyrics, as students can then be exposed to orthography later, with the transmediation from one sign system into another (Olshansky, 2018). Not relying heavily on words could also become an inclusive strategy for teachers to implement with students with dyslexia. A similar strategy was implemented for Maths, through the Spatial-Temporal Math (ST Math) program, a game-based software to learn secondary-level students' mathematics comprehension and proficiency through visual learning (Wendt et al, 2018). The results of such report showed that the schools implementing an ST Math program significantly outperformed control groups, thus supporting the idea that a less verbocentric educational system is a feasible option.

Moreover, extensive listening without comprehension as target should be encouraged, as students could also keep listening to such songs in their own time, which for teachers would mean making use of out-of-class exposure to their advantage. Such resolution would also allow for students to have a more guided exposure to authentic materials, since, although learning can occur both directly and indirectly, the direct mode is necessarily limited in its nature. As reported in Murphey (1992, p.6) "words listening first enables us to imitate, but not necessarily to understand immediately", so although students might not understand what they are listening to, this does not mean that such activity is to be considered pointless.

9486 words

34

Bibliography

Ajmi, A. (2016). Assessment: Anything, Anywhere, Anytime Using JotForm, Workshop at University of Missouri-Kansas City, 4-8-2016.

Albers, P. (2007). Finding the artist within: Creating and reading visual texts in the English language arts classroom. Newark, DE: International Reading Association.

Arnold, J.L. & Herrick, E. (2017). New Ways in Teaching With Music Paperback, TESOL Press, Annapolis Junction.

Beck, I. L., McKeown, M. G., & Kucan, L. (2002). Bringing words to life: Robust vocabulary instruction. New York, NY: Guilford Press.

Beall, M.L., Gill-Rosier J., Tate, J. Matten, A. (2008), State of the Context: Listening in Education, International Journal of Listening, v.22, p. 123-132 doi: org/10.1080/10904010802174826

Boersma, M. (2002). Praat, a system for doing phonetics by computer, *Glot International*, v.5 (9/10) p.341-345.

Brown, Ronan; Waring, Rob; Donkaewbua, Sangrawee. (2008) Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, v20 n2 p136-163.

Bongaerts, T., van Summerin, C., Planken, B., & Schils, E. (1997). *Age and ultimate attainment in the pronunciation of a foreign language*. Studies in Second Language Acquisition, 19, 447–465.

Campbell, R. (2011), The Power of the Listening Ear, *English Journal*, v.100 n5 p. 66-70

Chang, A.C.-S. (2009), Gains to L2 listeners from reading while listening vs. listening only in comprehending short stories, *System, 37 (4), 652-663*.

Chang, C. -S. A. (2012). Gains to L2 Learners from extensive listening: Listening development, vocabulary acquisition and perceptions of the intervention, *Hong Kong Journal of Applied Linguistics*, 14(1), 25-47.

Chang, C.-S., Read, J., (2006). The effects of listening support on the listening performance of EFL learners. *TESOL Quarterly 40, 375–397*.

Chung, W. L. (2008). The effectiveness of explicit, implicit, and noticing instruction: Mandarin speakers' perceptions and production of English sentence stress. PhD Dissertation, Alliant International University, San Francisco, CA, USA.

Christiner, M and Reiterer S.M. (2015), A Mozart is not a Pavarotti: singers outperform instrumentalists on foreign accent imitation, Frontiers in Human Neuroscience, v9 p482, doi: 10.3389/fnhum.2015.00482.

Copland, F., Garton, S., Burns, A. (2014). Challenges in Teaching English to Young Learners: Global Perspectives and Local Realities, *TESOL Quarterly*, doi: 10.1002/tesq.148.

Costenaro, V. (2011), Phonological Awareness and Reading Acquisition: An Educational Proposal for Introducing English in Italian Preschools, PhD

Dissertation, University of Padua, Italy doi: http://paduaresearch.cab.unipd.it/3268/.

Coumel M, Christiner M and Reiterer SM (2019) Second Language Accent Faking Ability Depends on Musical Abilities, Not on Working Memory. Front. Psychol. 10:257. doi: 10.3389/fpsyg.2019.00257.

Darcy, I. (2018), Powerful and Effective Pronunciation Instruction: How Can We Achieve It?, *CATESOL Journal*, v.30(1), p.13-45.

de Souza, R.A. & Soares Silva, J., (2015). Exploring the measurement of vocabulary size to differentiate Brazilian Portuguese-English bilinguals' access to grammatical knowledge in the L2, *Revista Linguíʃtica,* v. 11, n. 1 (2015).

de Vos, J., Schriefers, H., Mivard, M. and K. Lemhöfer. 2018. 'A meta-analysis and meta- regression of incidental second language word learning from spoken input,' Language Learning 68/4: 906−941.

Derwing, T. M. (2003). What do ESL students say about their accents? *Canadian Modern Language Review, 59, 545–564.*

Derwing T. M. and Munro, M.J. (2005), Second Language Accent and Pronunciation Teaching: A Research-Based Approach, *Tesol Quarterly*. doi: 10.2307/3588486

Dolean, D. D. (2015). The effects of teaching songs during foreign language classes on students' foreign language anxiety. *Language Teaching Research*, v.20, p.638–653. doi: 10.1177/1362168815606151.

Ellis, R. (2015), Understanding Second Language Acquistion, Oxford University Press.

Farina, M.A. (2013). The Effect Of Input Modality On Pronunciation Accuracy Of English Language Learners, PhD Dissertation, University of Central Florida, FL, USA doi: http://stars.library.ucf.edu/etd/2622

Failoni, J. W. (1993), Music as Means To Enhance Cultural Awareness and Literacy in the Foreign Language Classroom, *Mid-Atlantic Journal of Foreign Language Pedagogy*, v.1 p.97-108.

Flege, J. E. (1995) Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.) Speech perception and linguistic experience: Issue in cross-language research. Timonium, MD: York Press. p. 229-273.

Flege, J. E. (2009). Give input a chance! In T. Piske & M. Young-Scholten (Eds.), Input matters in SLA (pp. 175–218).

Fullana, N. (2006). *The development of English (FL) perception and production skills: Starting age and exposure effects*. In C. Muñoz (Ed.), Age and the rate of foreign language learning (pp. 41–64). Clevedon, UK: Multilingual Matters.

Garnier, M., & Schmidt, N. (2015). The PHaVE List: A pedagogical list of phrasal verbs and their most frequent meaning senses. *Language Teaching Research*, 19(6), 645-666.

Gass, S. and Selinker, L. (2001) *Second Language Acquisition: An Introductory Course*, New Jersey: Lawrence Erlbaum Associates Publishers, 2001:200.

Gilbert, J. (1995). *Pronunciation practice as an aid to listening comprehension.* In D. J. Mendelsohn and J. Rubin (Eds.), A guide for the teaching of second language listening. (pp. 97-102). San Diego, California: Dominie Press, Inc.

Goodwin, J. (2014). Teaching pronunciation. In M. Celce-Murcia, D. M. Brinton, & M. A. Snow (Eds.), *Teaching English as a second or foreign language* (4th ed., pp. 136-152). Boston, MA: National Geographic Learning/Heinle Cengage Learning.

Hindman, A. H., & Wasik, B. A. (2013). Vocabulary learning in Head Start: Nature and extent of classroom instruction and its contributions to children's learning. *Journal of School Psychology*, 51(3), 387-405.

Hulstijn, J.H. (2013). Incidental learning in second language acquisition. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 2632-2637). Malden, MA: Blackwell Publishing Ltd

Hulstijn, J.H. (2001), Intentional and incidental second-language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity, in *Cognition and Second Language Instruction,* p. 258-286, Cambridge University Press.

Isaacs, T. (2009). *Integrating form and meaning in L2 pronunciation instruction.* TESL Canada Journal/ Revue TESL Du Canada, 27, 1–12.

Jones, D. M. (1999). The cognitive psychology of auditory distraction. *British Journal of Psychology,* v.90, p.167-187

Kerekes, E. (2015). Using songs and lyrics in out-of-class learning. In D. Nunan & J. Richards (Eds.), Language learning beyond the classroom (pp. 33-42). New York: Routledge.

Kissling, E.M. (2013), Teaching Pronunciation: Is Explicit Phonetics Instruction Beneficial for FL Learners? *The Modern Language Journal*, 97, 3 doi: 10.1111/j.1540-4781.2013.12029.x

Krashen, S. D. (1982). *Principles and practice in second language acquisition.* Oxford: Pergamon.

Laufer, B. (2006). Comparing focus on form and focus on forms in second language vocabulary learning, *Canadian Modern Language Review*, 63(1), 149–166.

Leland, C., & Harste, J. (1994). Multiple ways of knowing: Curriculum in a new key. Language Arts, 71, 337–345.

Loewen, S. (2015) Introduction to Instructed Second Language Acquisition, New York, Routledge.

Mairano, P., Bassetti, B., Sokolovic-Perovic, M. and Cerni, T. (2018) Effects of L1 orthography and L1 phonology on L2 English pronunciation. *Revue Française de Linguistique Appliquée*, XXIII (1). pp. 45-57.

Maneshi, N. (2017). Incidental vocabulary learning through listening to songs PhD Dissertation. University of Western Ontario, Canada.

Mareschal, C., (2007). Student Perceptions of a Self-Regulatory Approach to Second Language Listening Comprehension Development. PhD Dissertation, University of Ottawa, Canada.

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

Miranda, D. (2013). The role of music in adolescent development: much more than the same old song, *International Journal of Adolescence and Youth,* v.18 p.5-22. doi: 10.1080/02673843.2011.650182

Murphey, T. (1990). The song stuck in my head phenomenon: A melodic din in the head? *System*, v.18, p.53–64. doi: 10.1016/0346-251X(90)90028-4.

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

Murphey, T. and Alber, J.L. (1985). A Pop Song Register: The Motherese of Adolescents as Affective Foreigner Talk, *TESOL Quarterly*, doi: 10.2307/3586679

Nation, I.S.P. (1990) Teaching and Learning Vocabulary New York: Newbury House.

Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. Language Learning, 50, 417-528.

Olshansky, B. (2018), The universal language of pictures: A critical tool for advancing student writing, TESOL Journal, doi: org/10.1002/tesj.402

Pavia, N., Webb S., Farahnaz Faez, F. (2019), Incidental Vocabulary Learning through Listening to Songs, *Studies in Second Language Acquisition*, p.1-24. doi:10.1017/S0272263119000020

Pellicer-Sánchez, A. (2018). Examining second language vocabulary growth: Replications of Schmitt (1998) and Webb & Chang (2012), *Language Teaching*, doi: 10.1017/s026144481800037x

Pellicer-Sánchez, A. & Norbert Schmitt, (2010). Incidental vocabulary acquisition from an authentic novel : Do Things Fall Apart ? *Reading in a Foreign Language,* v22 n1 p31-55 Apr 2010

Peters A.E., Noreillie, S., Heylen K., Bulté, B., Desmet P. (2019). The Impact of Instruction and Out-of-School Exposure to Foreign Language Input on Learners' Vocabulary Knowledge in Two Languages, *Language Learning*, doi: 10.1111/lang.12351

Rahman, M. (2007). Development of Language through Music in EFL Classroom, AlQattan Centre for Educational Research and Development. Retrieved from: www.qattanfoundation.org/pdf/1401_2.doc

Rallo Fabra, L. (2016). Perception and production of L2 speech sounds: L1 phonetic/phonological categories do not tell it all in Fernández Planas, A. Ma. (ed.) Reflexiones sobre aspectos de la fonética y otros temas de lingüística, Barcelona, p. 417-425.

Richards, H & Durrant, C. (2003). To Sing or Not to Sing: A study on the development of `non-singers' in choral activity, *Research Studies in Music Education*,

Rundell, M. (2002). Macmillan English Dictionary: For Advanced Learners, Macmillan Education, London. doi: 10.1177/1321103X030200010501

Saito, K., Suna, H. and Tierneyb, A. (2019), Explicit and Implicit Aptitude Effects on Second Language Speech Learning: Scrutinizing Segmental and Suprasegmental Sensitivity and Performance via Behavioural and Neurophysiological Measures, *Bilingualism: Language and Cognition*, doi: 10.1017/S1366728918000895.

Schmidt, R. (1990). The Role of Consciousness in Second Language Learning. *Applied Linguistics*, v.11, p.129-158.

Schwarz, M. (2013). Learning with Lady GaGa & CO: Incidental EFL vocabulary acquisition from pop songs. *Views*, v.22, p.17–48.

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(2). doi: 10.1002/tesq.445.

Singleton, D. (1999). Exploring the second language mental lexicon. Cambridge: Cambridge University Press.

Staehr, L. S. (2008). Vocabulary Size and the Skills of Listening, Reading and Writing, *Language Learning Journal*, v.36, p.139-152.

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.

Tominaga, Y. (2009), An Analysis of Successful Pronunciation Learners: In Search of Effective Factors in Pronunciation Teaching, *Journal of Pan-Pacific Association of Applied Linguistics*, v13 n1 p127-140 2009

Vandergrift, L., 2007. Recent developments in second and foreign language listening comprehension research. Language Teaching 40, 191–210.

Villalobos Ulate, M. (2007), Using songs to improve EFL students' pronunciation, I International Applied Linguistics Congress - Universidad Nacional de Costa Rica. doi

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.

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

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.

Wendt, S., Rice, J., Nakamoto J. (2018), MIND Research Institute's ST Math Program and Math Performance, Retrieved from: https://www.mindresearch.org/our-research

Wheelock, A. (2016). Phonological Difficulties Encountered by Italian Learners of English : An Error Analysis, *Hawaii Pacific University TESOL Working Paper Series 14, 41-61*.

Appendix A: Lesson Plans

Lesson Plan for explicit teaching of pronunciation

Time	Activity
5'	watch video of song on Youtube without lyrics. (listening 1)
5'	ask general questions (i.e. do you like this song? Why/Why not?)
5'	listen to the song and follow the lyrics/no lyrics worksheet (5') (listening 2)
5'	ask them which words were singled out (5') – what do they have in common?
10'	explanation and drills with difference on how I/we pronounce /th/ or /h/ and how the singer pronounces it.
15'	Pictionary with Chinese Whispers with Pictionary (in teams – the pronunciation needs to be correct – peer assessment for corrections)
5'	Plenary: listen only and sing along (listening 4)

Lesson Plan for explicit teaching of vocabulary

Time	Activity
5'	watch video of song on Youtube without lyrics. (listening 1)
5'	ask general questions on singer (i.e. do you know him? do you know his name?)
5'	listen to the song and follow the lyrics/no lyrics worksheet (5') (listening 2)
5'	ask them which words were singled out (5') – what do they have in common?
10'	Have a slide with images, explain to the them what the words mean and then have drills for practice
15'	Hot Seat in groups to ask for meaning of words (song in the background while they prepare questions– listening 3)
5'	Plenary: listen only and sing along (listening 4)

Appendix B: Lyrics

Lyrics for Move Along – The All American Rejects

Go **ahead** as you waste your days with thinking
When you fall, everyone stands
Another day, and you've **had** your fill of sinking
With the life **held** in your
Hands are shaking cold
These hands are meant to **hold**
Speak to me
When all you got to keep is strong
Move along, move along like I know you do
And even when your **hope** is gone
Move along, move along just to make it through
Move along
Move along
So a day when you've lost yourself completely
Could be a night when your life ends
Such a **heart** that will lead you to deceiving
All the pain **held** in your
Hands are shaking cold
Your **hands** are mine to **hold**
Speak to me
When all you got to keep is strong
Move along, move along like I know you do
And even when your **hope** is gone
Move along, move along just to make it through
Move along
(Go on, go on, go on, go on)
When everything is wrong, we move along
(Go on, go on, go on, go on)
When everything is wrong, we move along
Along, along, along, along
When all you got to keep is strong
Move along, move along like I know you do
And even when your **hope** is gone
Move along, move along just to make it through
When all you got to keep is strong
Move along, move along like I know you do
And even when your **hope** is gone
Move along, move along just to make it through
When all you got to keep is strong
Move along, move along like I know you do (Know you do)
And even when your **hope** is gone
Move along, move along just to make it through
Right back what is wrong
We move along
(Go on, go on, go on, go on)
Right back what is wrong
We move along
(Go on, go on, go on, go on)
Right back what is wrong
We move along

(Go on, go on, go on, go on)
Right back what is wrong
We move along

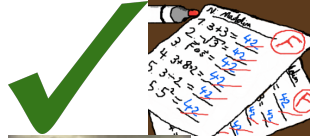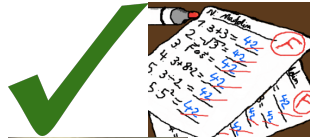## Lyrics for Have a Nice Day – John Bon Jovi

Why, you wanna tell me how to live my life?
Who, are you to tell me if it's black or white?
Mama, can you hear me? try to understand
Is innocence the difference between a boy and a man
My daddy lived the lie, it's just the price that he paid
Sacrificed his life, just slavin' away
Oh, if there's one thing I hang onto
That gets me through the night
I ain't gonna do what I don't want to
I'm gonna live my life
Shining like a diamond, rolling with the **dice**
Standing on the **ledge**, I show the wind how to fly
When the world gets in my face
I say, have a nice day
Have a nice day
Take a look around you; nothing's what it seems
We're living in the broken home of hopes and dreams
Let me be the first to shake a helping hand
Anybody **brave** enough to **take a stand**
I've knocked on every door, on every **dead end street**
Looking for forgiveness
What's left to believe?
Oh, if there's one thing I hang onto
That gets me through the night
I ain't gonna do what I don't want to
I'm gonna live my life
Shining like a diamond, rolling with the **dice**
Standing on the **ledge**, I show the wind how to fly
When the world gets in my face
I say, have a nice day
Have a nice day
Oh, if there's one thing I hang onto
That gets me through the night
I ain't gonna do what I don't want to
I'm gonna live my life
Shining like a diamond, rolling with the **dice**
Standing on the **ledge**, I show the wind how to fly
When the world gets in my face
I say, have a nice day
Have a nice day
Have a nice day
Have a nice day
Have a nice day
When the world keeps trying, to drag me down
I've gotta raise my hands, gonna **stand my ground**
Well I say, have a nice day
Have a nice day
Have a nice day

Appendix C: No-Lyrics
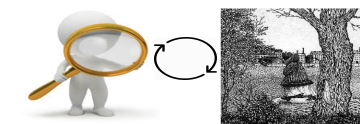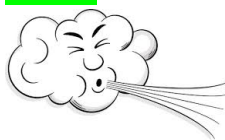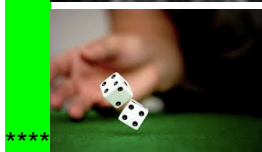
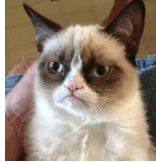## No Lyrics for Move Along – The All American Rejects
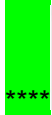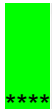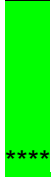
No Lyrics for Have a Nice Day – John Bon Jovi

****

****

****

****

****

****
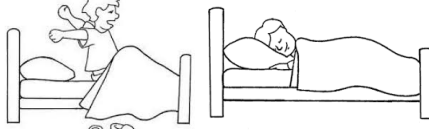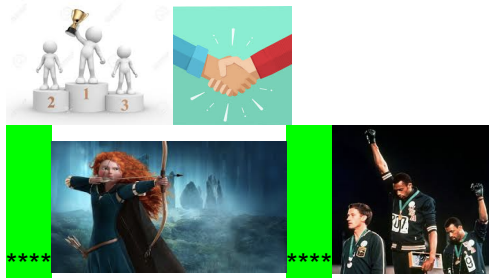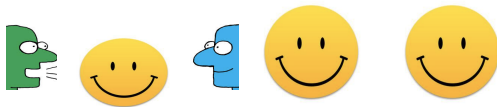
Appendix D: Production Tests

Production pre-test and post-test

Pronunciation Production Test
Registra la tua voce premendo RECORD mentre leggi le parole qui riportate.

(Record your voice clicking on RECORD while you read the words below)

1. hands
2. ahead
3. had
4. held
5. hold
6. hope
7. heart
8. the
9. there
10. thing
11. through
12. that
13. with
14. nothing

Vocabulary Production Test

Completa le frasi con le parole mancanti

(Complete the sentences with the missing words)

15. Don´t was_____ my time.

16. Don't run on the stairs, you might fa_____.

17. The police told the bystanders to mo_____ alo_____.

18. I have had my fi_____ of emotions for today.

19. The ship is sin _____.

20. I know this is difficult, but we can ma_____ i_____ thr_____.

21. He is always telling lies. He's a very dece_____ person.

22. You need to ri_____what you did wrong. Stop fighting.

23. It was a gorgeous day; the sun was shi_____ bright.

24. You should all be sta_____ when the teacher comes in.

25. To play monopoli you need a di_____.

26. The dove is eating on the window led_____.

27. I had to sta_____ m_____ gro_____ , I could not do what he wanted
    me to.

28. Firemen are very br_____.

29. Let's ta_____ a sta_____! This is unfair!

30. Make a U-turn! This is a de_____ e_____ st_____!

Appendix E: Perception/Reception Tests

Perception/Reception pre-test, immediate and delayed post-test

Pronunciation Perception Test
Ascolta la registrazione e decidi se le due parole sono uguali o diverse

(*Listen to the recording and decide whether the two words you hear are the same or different*)

1. ☐ same
   ☐ different

2. ☐ same
   ☐ different

3. ☐ same
   ☐ different

4. ☐ same
   ☐ different

5. ☐ same
   ☐ different

6. ☐ same
   ☐ different

7. ☐ same
   ☐ different

8. ☐ same
   ☐ different

9. ☐ same
   ☐ different

10. ☐ same
    ☐ different

11. ☐ same
    ☐ different

12. ☐ same
    ☐ different

13. ☐ same
    ☐ different

14. ☐ same
    ☐ different

50

Vocabulary Reception
Scegli quale delle opzioni è la definizione della parola in maiuscolo

(Choose which option is the definition of the word in capital letters)

15. TO WASTE

- o to use more of something than is necessary, or to use it in a way that does not produce the best results.
- o to have a particular flavour.
- o to glue paper onto a surface using paste.
- o to get more and more of something over a period of time.
- o I don't know

16. TO FALL

- o to move upwards or to a higher position.
- o to be in a particular state as a result of an emotion or a physical feeling.
- o to move quickly downwards from a higher position, usually by accident.
- o to move more slowly than other people so that you are behind them.
- o I don't know

17. TO MOVE ALONG

- o to try to prevent something from happening, especially because you do not approve of it or think it is harmful.
- o to do something so difficult that it seems almost impossible.
- o to prepare to destroy or defeat someone or something.
- o to progress or develop, or to make something progress or develop.
- o I don't know

18. FILL

- o someone who does another person's work while they are away.
- o an act of filling something until it is completely full, especially the petrol tank of a car.
- o when you feel no emotion, interest, or purpose.
- o so much of something that you do not want any more.
- o I don't know

19. SINKING

- o becoming completely understood.
- o disappearing below the surface of the water.

- o succeeding or failing, without any help.
- o disappearing above the surface of the water.
- o I don't know

20. MAKE IT THROUGH

- o to die as a result of an illness or an accident.
- o not to tell the truth.
- o to treat something as if it is more important than it really is.
- o to survive a dificult time.
- o I don't know

21. DECEIVING

- o tricking someone by behaving in a dishonest way.
- o producing the final result of a situation or event.
- o approving of an idea or of a person or organization and help them to be successful.
- o helping someone when they are having a dificult time.
- o I don't know

22. TO RIGHT

- o to make a right turn at a crossroads.
- o to be right in a discussion.
- o to have the right to do something.
- o to make something go back into the right state.
- o I don't know

23. SHINING

- o scaring.
- o bright.
- o with no light.
- o being shy.
- o I don't know.

24. STANDING

- o having your body in a downright position not supported by your feet.
- o holding something firmly without shaking or moving it.
- o having your body in an upright position supported by your feet.
- o holding something firmly shaking and moving it.
- o I don't know

25.DICE

- o a flat piece of food that has been cut from something larger
- o water that has frozen and become solid
- o a rectangular piece of paper with spots or images to play
- o a small block of wood or plastic with six sides marked with spots,
- o I don't know

26. LEDGE

- o the part of something that is furthest from its centre.
- o a narrow surface that continues out from the side of a cliff, wall, or other surface.
- o a piece of equipment consisting of a metal frame with a strong material stretched across it that you can jump up and down on for exercise or as a sport.
- o a line of bushes or small trees growing close together around a garden or field.
- o I don't know

27. TO STAND MY GROUND

- o not retreat in the face of opposition.
- o to stand up from the ground.
- o to mow the lawn of my garden.
- o retreat in the face of opposition.
- o I don't know

28. BRAVE

- o cautious.
- o capable of dealing with danger or pain, seeming to be frightened.
- o capable of dealing with danger or pain, without seeming to be frightened.
- o incapable of dealing with danger or pain, seeming to be frightened.
- o I don't know

29. TO TAKE A STAND

o   to stand up.

o   to state your opinion and refuse to change it.

o   to state your opinion and then change it.

o   to go to a fair.

o   I don't know

30. DEAD-END STREET

o   a dangerous street in a city

o   an end of a something that has multiple exits.

o   a street in an open neighbourhood.

o   an end of something that has no exit.

o   I don't know

Appendix F: Results

Main effects and interactions

Target	Source	F	df1	df2	Sig.
Pronunciation Production	Time	.106	1	166	.745
	Modality	6.504	1	166	.012*
	Instruction	.517	1	166	.473
	Modality*Time	.147	1	166	.702
	Instruction*Time	3.074	1	166	.081
	Modality*Instruction	Unable to compute due to 0 degrees of freedom			
	Modality*Instruction*Time	Unable to compute due to 0 degrees of freedom			

Post-hoc comparisons

	Time			Modality			Degree of Directness		
	T2-T1	.012	.927	No Lyrics – Lyrics	.346	.026*	Direct – Indirect	-.083	.536

Modality*Time

		Time	
Lyrics	T2-T1	.141	.377
No Lyrics	T2-T1	-.274	.223

Instruction*Time

		Time	
Indirect	T2-T1	-.198	.230
Direct	T2-T1	.421	.072

Modality*Instruction

		Modality	
		No Lyrics –Lyrics	
No Lyrics –Lyrics	Indirect	.401	.028*
	Direct	n/a	n/a

Modality* Instruction *Time

			Time	
Lyrics	Indirect	T2-T1	-.130	.574
No Lyrics	Indirect	T2-T1	-.274	.223
Lyrics	Direct	T2-T1	.421	.072
No Lyrics	Direct	T2-T1	n/a	n/a

Target	Source	F	df1	df2	Sig.
Pronunciation / Perception	Time	4.804	2	1,794	.008**
	Modality	12.313	1	1,794	.000***
	Instruction	.010	1	1,794	.921
	Modality*Time	2.389	2	1,794	.092
	Instruction*Time	.955	2	1,794	.385
	Modality*Instruction	.740	1	1,794	.390
	Modality*Instruction*Time	.606	3	1,794	.611

Time (pairwise comparisons)

	Comparison	Mean Diff	Sig.
Time	T2-T1	.059	.020*
	T3-T1	.069	.026*
	T3-T2	-.000	.986
Modality*Time — No Lyrics	T2-T1	.053	.119
	T3-T1	.085	.021*
	T3-T2	.032	.245
Modality*Time — Lyrics	T2-T1	.067	.042*
	T3-T1	.017	.603
	T3-T2	-.049	.124
Instruction*Time — Direct	T2-T1	.079	.025*
	T3-T1	.091	.016*
	T3-T2	-.013	.673
Instruction*Time — Indirect	T2-T1	.041	.210
	T3-T1	.027	.421
	T3-T2	.011	.681
Modality*Instruction*Time — Lyrics / Indirect	T2-T1	.070	.118
	T3-T1	-.000	1.000
	T3-T2	-.070	.135
No Lyrics / Indirect	T2-T1	.017	.686
	T3-T1	.042	.335
	T3-T2	.024	.546
Lyrics / Direct	T2-T1	.064	.145
	T3-T1	.033	.473
	T3-T2	-.031	.469
No Lyrics / Direct	T2-T1	.095	.060
	T3-T1	.134	.013*
	T3-T2	.040	.278

Modality (pairwise comparisons)

	Comparison	Mean Diff	Sig.
Modality	No Lyrics – Lyrics	.063	.006**
Modality*Instruction — Indirect	No Lyrics – Lyrics	.082	.014*
Modality*Instruction — Direct	No Lyrics – Lyrics	.044	.115

Degree of Directness (pairwise comparisons)

	Comparison	Mean Diff	Sig.
Degree of Directness	Direct – Indirect	.002	.921

56

Target: Vocabulary Production

Source	F	df1	df2	Sig.	Time				Modality				Degree of Directness			
Time	239.901	1	1,400	.000***		T2-T1	.391	.000***		No Lyrics – Lyrics	.0039	.129		Direct – Indirect	.041	.139
Modality	2.507	1	1,400	.114												
Instruction	2.338	1	1,400	.126												
Modality*Time	0.409	1	1,400	.523												
Instruction*Time	14.731	1	1,400	.000***	Lyrics	T2-T1	.361	.000**	T1	No Lyrics = Lyrics	.009	.590	T1	Direct = Indirect	-.020	.223
					No Lyrics	T2-T1	.432	.000**	T2	No Lyrics = Lyrics	.089	.033*	T2	Direct = Indirect	.219	.000***
					Indirect	T2-T1	.273	.000**								
					Direct	T2-T1	.512	.000**								
Modality*Instruction	4.220	1	1,400	.040*					No Lyrics = Lyrics	Indirect	.090	.014*				
									No Lyrics = Direct		-.022	.608				
Modality*Instruction*Time	3.606	2	1,400	.017*	Lyrics Indirect	T2-T1	.259	.000**	T1 Indirect	No Lyrics = Lyrics	.063	.030*	T1 Lyrics	Direct = Indirect	.023	.227
					No Lyrics Indirect	T2-T1	.277	.000**	T2 Indirect	No Lyrics = Lyrics	.081	.181	T2 Lyrics	Direct = Indirect	.213	.000***
					Lyrics Direct	T2-T1	.449	.000**	T1 Direct	No Lyrics = Lyrics	-.031	.139	T1 No Lyrics	Direct = Indirect	-.071	.024*
					No Lyrics Direct	T2-T1	.569	.000**	T2 Explici.	No Lyrics = Lyrics	.089	.147	T2 No Lyrics	Direct = Indirect	.221	.003**

Target: Vocabulary Reception

Source	F	df1	df2	Sig.
Time	34.837	2	2,100	.000***
Modality	5.744	1	2,100	.017*
Instruction	16.803	1	2,100	.000***
Modality*Time	1.424	2	2,100	.241
Instruction*Time	2.012	2	2,100	.134
Modality*Instruction	.009	1	2,100	.908
Modality*Instruction*Time	.112	3	2,100	.857

Time (pairwise comparisons)

Time (main):
Time	Mean diff	Sig.
T2-T1	.214	.000***
T3-T1	.165	.000***
T3-T2	-.049	.078

Modality*Time – Lyrics:
Time	Mean diff	Sig.
T2-T1	.186	.000***
T3-T1	.116	.000***
T3-T2	-.070	.041

Modality*Time – No Lyrics:
Time	Mean diff	Sig.
T2-T1	.242	.000***
T3-T1	.215	.000***
T3-T2	-.027	.542

Instruction*Time – Indirect:
Time	Mean diff	Sig.
T2-T1	.151	.000***
T3-T1	.122	.001**
T3-T2	-.029	.445

Instruction*Time – Direct:
Time	Mean diff	Sig.
T2-T1	.279	.000***
T3-T1	.210	.000***
T3-T2	-.069	.080

Modality*Instruction*Time – Lyrics / Indirect:
Time	Mean diff	Sig.
T2-T1	.113	.010*
T3-T1	.065	.122
T3-T2	-.047	.296

Modality*Instruction*Time – No Lyrics / Indirect:
Time	Mean diff	Sig.
T2-T1	.190	.001**
T3-T1	.182	.001**
T3-T2	-.008	.899

Modality*Instruction*Time – Lyrics / Direct:
Time	Mean diff	Sig.
T2-T1	.264	.000***
T3-T1	.171	.000***
T3-T2	-.092	.060

Modality*Instruction*Time – No Lyrics / Direct:
Time	Mean diff	Sig.
T2-T1	.294	.000***
T3-T1	.249	.000***
T3-T2	-.046	.457

Modality (No Lyrics – Lyrics)

	Mean diff	Sig.
No Lyrics – Lyrics	.053	.018*
T1: No Lyrics – Lyrics	.004	.903
T2: No Lyrics – Lyrics	.060	.129
T3: No Lyrics – Lyrics	.103	.008**

No Lyrics – Lyrics	Direct	Indirect
Direct	.055	.050
Indirect	.074	.120

Degree of Directness (Direct – Indirect)

	Mean diff	Sig.
Direct – Indirect	.101	.000***
T1 Direct – Indirect	.031	.360
T2 Direct – Indirect	.159	.000***
T3 Direct – Indirect	.119	.003**
T1 Lyrics Direct – Indirect	.019	.642
T2 Lyrics Direct – Indirect	.170	.000***
T3 Lyrics Direct – Indirect	.125	.008**
T1 No Lyrics Direct – Indirect	.043	.436
T2 No Lyrics Direct – Indirect	.148	.027*
T3 No Lyrics Direct – Indirect	.110	.103