INVESTIGATING SONG-BASED LANGUAGE TEACHING AND ITS EFFECT ON LEXICAL LEARNING

by

FRIEDERIKE ANNA GERDA TEGGE

A thesis submitted to the Victoria University of Wellington in fulfilment of the requirements for the degree of Doctor of Philosophy

Victoria University of Wellington

Abstract

The present thesis addresses the following question: Can popular songs as they are currently used in second and foreign language classrooms benefit lexical learning? Lexical learning is defined as the acquisition of new vocabulary as well as the consolidation and further elaboration of familiar words and phrases. To answer this research question, three methodologically distinct studies are reported.

In the first study an international questionnaire explored teacher cognitions as well as actual teaching practices involving songs. The responses of 568 informants in 41 countries indicate that a majority of respondents believe in the usefulness of songs for language learning and that many respondents utilize songs in class for clearly defined pedagogical purposes, including vocabulary learning. The questionnaire also elicited information from the respondents about the way they incorporate songs in lessons, including details about how often a song is played and what types of form- and meaning-focused activities are used to engage learners with the lyrics of a song.

The second study investigated the lexical characteristics of teacher-selected songs and the vocabulary learning opportunities they afford. For this purpose, a corpus of 635 songs used for ESL/EFL purposes, comprising 177,384 tokens, was compiled and analysed. Results indicate that teacher-selected songs are short, repetitive and relatively undemanding as far as lexis is concerned compared to other authentic text genres. Knowledge of the 4000 most frequent word families of English provides 98% coverage of the running words in this song corpus. Little difference was found in terms of the overall vocabulary demands between songs intended for use with beginners, intermediate and advanced learners.

The third study investigates whether participating in a song-based lesson results in higher verbatim text retention compared to a lesson based on a poem or a prose text. For the sake of ecological validity, the procedures and the materials used in the classroom intervention study were informed by the findings of the teacher questionnaire (study one) and the song corpus analysis (study two). Results indicate that a song-based

i

language lesson but also a poem-based lesson result in significantly higher recognition and cued recall of verbatim text than a lesson based on a prose text.

In response to the overall question, this thesis provides evidence that songs as they currently tend to be used by language teachers around the world indeed benefit certain aspects of lexical learning, perhaps in particular the entrenchment in memory of already (half-)familiar words in association with their phraseological patterning. It is argued that, while certain structural characteristics of songs (and poems) have the potential of rendering text (and the lexis therein) memorable, it is the way that songs tend to be exploited in the classroom that capitalizes on this mnemonic potential.

Acknowledgements

I would like to express my sincerest gratitude to my doctoral supervisors Professor Frank Boers and Dr Averil Coxhead. Thank you for your guidance, insight, and encouragement, your always open doors and creative ways to help me when I was "lost in the woods". I am also grateful to everyone in the School of Linguistics and Applied Language Studies (LALS) for all your help, input and the incredibly positive atmosphere.

I am grateful to all my survey informants, who generously shared their experiences, knowledge, ideas and questions with me. Thank you also to all my experimental participants. It was truly a pleasure to work with you. I am particularly grateful to the institutions and teachers, who supported this research by giving me access to their students and by assisting me in many ways during data collection. In particular, I want to thank Shirley Elsas, Bente Opel, Christian Spickermann, Dr Jelena Vujić, Dr Mirjana Daničić and Dr Biljana Cubrović.

A big thank you to my family and friends in Germany, New Zealand and around the world, who encouraged me, hosted me, fed me ... In particular, I want to thank my mom for oh so many things and above all for being there through it all. Danke Muddi! I also want to express my deep gratitude to Audrey and Peter Vernon. Your friendship has been invaluable to me. Thank you to my "partners in crime" Simin Littschwager and Jackie Yeoh, and to Annesley and Monty, the best flatmates I could have wished for. I would also like to thank my friends at Central Baptist and particularly my family group and the Macs. Thank you also to all my awesome office-mates in legendary VZ410. Thanks for all the good times, for wearing stripes on "show-your-office-spirit" day, for helping me out when I hobbled around on crutches and for tolerating the crazy during the final weeks of writing.

I am indebted to professors, lecturers and teachers who inspired me throughout my "academic life". Particularly, I would like to thank Professor Stayc DuBravac, Dr Prisca Augustyn, Professor Willis Edmondson, Dr Johannes Eckerth and Helen Stein. Finally, I also want to express my gratitude to Victoria University of Wellington for the financial support, without which I would not have been able to conduct this research.

iii

Table of Contents

Abstract	i
Acknowledgements	iii
Table of Contents	v
List of Tables	ix
List of Figures	xii
List of Appendices	xiii
Chapter 1 Introduction	1
1.1 Introduction: Why study the use of pop songs in the language classroom?	1
1.2 Why study the learning of vocabulary through songs?	3
1.3 Why is vocabulary learning relevant and what does it entail?	4
1.4 Why is it relevant to remember verbatim text and learn multi-word units?	5
1.5 The aim and significance of the present study	6
1.6 Organisation of the thesis	7
Chapter 2 Literature review	9
2.0 Introduction	9
2.1 Study 1: Surveying teacher cognitions and classroom practices	10
2.1.1 Exploring teachers' beliefs and classroom practices	10
2.1.2 Implications of previous research for the current survey study	14
2.2 Study 2: The lexical profile of songs used in the language classroom	15
2.2.1 Existing research on the lexical profile of songs in language teaching	15
2.2.2 Methodological considerations: Assessing the lexical demand of a text genre	19
2.2.3 The vocabulary size necessary to comprehend various text genres	21
2.2.4 Can songs be considered written or spoken texts?	26
2.2.5 Implications from existing research for the present corpus study	28
2.3 Study 3: The benefits of songs for lexical learning	29
2.3.1 Pertinent models of information processing and storage	
2.3.2 Cognitive-psychological research on the potential mnemonic effect of songs an	d poems 36
2.3.3 The effect of melody, rhythm and lexical structural regularities on memory	40
2.3.4 Research on a mnemonic effect of songs in L2 studies	45
2.4 A summary of the literature review	53
Chapter 3 The teacher survey	55

3.0 Introduction	55
3.1 Research questions	56
3.2 Methodology	56
3.2.1 Survey design and piloting	57
3.2.2 Ethical considerations	58
3.2.3 Distribution	58
3.2.4 Informants	59
3.3 Data analysis	66
3.4 Results	66
3.4.1 The informants' views on songs' usefulness and their actual use of s	ongs in class 67
3.4.2 Teaching goals: For what purposes do informants use songs?	71
3.4.3 Informants' reasons for not using songs	74
3.4.4 Factors influencing the informants' use of songs	
3.4.5 A focus on vocabulary – beliefs and teaching practices	83
3.4.6 How are songs being used by informants?	
3.4.7 Song recommendations	
3.5 Discussion	
3.5.1 If and how songs are used in language classrooms	
3.5.2 Teachers' purposes of song use	
3.5.3 Teachers' use of songs to teach vocabulary	
3.5.4 Teachers' reasons for not using songs	
3.6 Limitations of the survey study	
3.7 Conclusion	
Chapter 4 The pedagogical song corpus study	103
4.0 Introduction	
4.1. Research questions	
4.2 Measuring lexical demand and vocabulary learning opportunities	
4.3 Methodology	105
4.3.1. The songs	106
4.3.2 The procurement and editing of the lyrics	108
4.4 Data analysis	110
4.5 Results	111
4.5.1 Common characteristics of songs used in ESL/EFL classrooms	111
4.5.2 The lexical profile of songs	

4.5.3 The vocabulary load of songs used with learners of different proficiency le	vels 120
4.5.4 Comparing the vocabulary load of songs with the lexical demand of other a	authentic
genres	
4.6 Discussion	
4.7 Limitations of the corpus study	
4.8 Conclusion	
Chapter 5 The quasi-experimental intervention study	132
5.0 Introduction	
5.1 How can episodic memory aid vocabulary learning?	
5.2 Research question	
5.3 Methodology	
5.3.1 Participants	
5.3.2 Materials used during the intervention	
5.3.3 The intervention – a typical language lesson	
5.3.4 The test instruments	
5.4 Data analysis and test scoring	
5.5 Pilot study	
5.5.1 Implications of the pilot study for the main study	
5.6 Results	
5.6.1 Results for each participating institution (Serbia, Germany, Belgium)	
5.6.2 Discussion of the overall results at three institutions	156
5.6.3 Results for each memory test (free recall, cued recall and recognition)	
5.6.4 Discussion of the separate results for each memory test	
5.7 Overall discussion of the findings of the intervention study	
5.7.1 Second language learners' retention of verbatim text	
5.7.2 The effects of a song-based and a prose-based treatment on verbatim text	retention
5.7.3 The effect of a poem-based intervention on verbatim retention	
5.7.4 Moderating variables: the participants' language proficiency and the test t	ype 168
5.7.5 The ecological validity of the intervention	
5.8 Limitations of the intervention study	
5.9 Conclusion	
Chapter 6 Overall discussion	175
6.0 Introduction	

Α	ppendix	217
R	eferences	198
	6.9 Future research	195
	6.8 Limitations of the thesis as a whole	193
	6.7 Original contributions of the study	192
	6.6 Pedagogical implications of the thesis	188
	6.5.3 Song-based teaching, cognitive elaboration and episodic memory	187
	6.5.2 The effect of the written line-by-line format	185
	6.5.1 Structural regularities and a greater attention to surface form	180
	6.5 Potential factors contributing to a mnemonic effect of songs and poems	178
	6.4 Can songs as they are currently used in second language classrooms benefit lexical learning?	178
	6.3 Study 3: The quasi-experimental intervention study	177
	6.2 Study 2: The pedagogical song corpus	176
	6.1 Study 1: The teacher survey	175

List of Tables

Table 2.1	Lexical knowledge, indicated in word families, required for 95% and 98% coverage of different authentic text genres	22
Table 3.1	Languages taught by the informants	60
Table 3.2	The informants' institution(s)	61
Table 3.3	The informants' nationalities	62
Table 3.4	The informants' geographical location	63
Table 3.5	The informants' teaching experience with particular proficiency levels	64
Table 3.6	The informants' teaching experience with particular class types	65
Table 3.7	Importance of music in private life	65
Table 3.8	Informants' listening habits (vocal music)	65
Table 3.9	Informants singing in a vocal ensemble or playing an instrument	66
Table 3.10	Usefulness of songs to foster language acquisition	67
Table 3.11	Usefulness of songs when teaching different proficiency levels	68
Table 3.12	Usefulness of songs when teaching different class types	68
Table 3.13	Use of songs with learners of different proficiency levels	70
Table 3.14	Use of songs with different class types	71
Table 3.15	Informants' purposes of using songs (prefabricated answer choices)	72
Table 3.16	Informants' purposes of using songs (open-ended answers)	73
Table 3.17	Responses to the question – why don't you use songs?	74
Table 3.18	Informants' use of songs and frequency of listening to vocal music	77
Table 3.19	Informants' country (by region) and the use of songs	79
Table 3.20	Informants' attitude towards song use by region	79
Table 3.21	Song use by institution	80
Table 3.22	Informants' attitude towards song use by gender	81

Table 3.23	Informants' song use by gender	81
Table 3.24	Gender ratio by region, compared to the use of songs	82
Table 3.25	Informants' opinion about the use of songs to introduce new vocabulary	83
Table 3.26	Informants' use of songs to introduce new vocabulary	84
Table 3.27	Informants' opinion about the use of songs to practice familiar vocabulary	84
Table 3.28	Informants' use of songs to practise familiar vocabulary	84
Table 3.29	Aspects rendering songs suitable for vocabulary learning (prefabricated answers)	85
Table 3.30	Aspects rendering songs suitable for vocabulary learning (open answers)	85
Table 3.31	Class-time spent on a song and complementary activities	88
Table 3.32	The number of sing-alongs	88
Table 3.33	The number of times informants repeat a song per lesson	89
Table 3.34	The replaying of a song in a subsequent lesson	89
Table 3.35	Activities used with songs in class (prefabricated answer options)	90
Table 3.36	The informants' description of their learners' response	90
Table 3.37	Problems experienced when using songs in class (prefabricated answers)	91
Table 3.38	Problems experienced when using songs in class (open-answer categories)	92
Table 4.1	The average number of words in the corpus songs	112
Table 4.2	The period of publication of the songs found in the corpus	112
Table 4.3	The country of origin of the singers	113
Table 4.4	The frequencies of words in the song corpus in tokens, types and families	116
Table 4.5	Cumulative coverage, without and with proper nouns, transparent compounds and marginal words	117

Table 4.6	Coverage in the three subcorpora (tokens only)	120
Table 4.7	Cumulative coverage for the complete corpus and the subcorpora including proper nouns, transparent compounds and marginal words	121
Table 4.8	Lexical knowledge required for 95% and 98% coverage of different authentic text genres	123
Table 4.9	Coverage (in percent) provided by high- to mid-frequency words in different authentic text genres	123
Table 5.1	Serbia – performance on three memory tests during immediate posttesting for three learning conditions	152
Table 5.2	Serbia – performance on three memory tests during delayed testing for two learning conditions	153
Table 5.3	Germany – performance on three memory tests during immediate testing for two learning conditions (song, prose)	153
Table 5.4	Germany – performance on three memory tests during delayed testing for two learning conditions (song, prose)	154
Table 5.5	Belgium – performance on three memory tests during immediate testing across two learning conditions (song, poem)	155
Table 5.6	Belgium – performance on three memory tests during delayed testing across two learning conditions (song, poem)	156
Table 5.7	Free recall: Descriptive statistics for Serbia, Germany and Belgium	160
Table 5.8	Free recall: Inferential statistics for Serbia, Germany, Belgium	160
Table 5.9	Cued recall: Descriptive statistics for Serbia, Germany and Belgium	161
Table 5.10	Cued recall: Inferential statistics for Serbia, Germany and Belgium	161
Table 5.11	Recognition: Descriptive statistics for Serbia, Germany and Belgium	162
Table 5.12	Recognition: Inferential statistics for Serbia, Germany and Belgium	162
Table 6.1	Retention of words in line-final position in immediate recall in Serbia	184

List of Figures

Figure 1.1	The interconnection of the three studies investigating three different aspects of song use (teaching practices, material, learning outcome)	8
Figure 3.1	An overview of the teacher survey	55
Figure 3.2	A breakdown of the questionnaire	57
Figure 4.1	Word knowledge (in 1,000-word frequency bands) required to reach 98% coverage of 108 songs	119
Figure 5.1	Overview of the quasi-experimental intervention and post-test battery	135
Figure 5.2	An example of the prose text with highlighted differences to the other formats	139
Figure 5.3	Description of the procedures of the intervention (including material and timing)	140
Figure 5.4	Expressions from the target text and their translations in the participants' native languages presented in a pre-teaching activity	143
Figure 5.5	Excerpt from the cued-recall test	147
Figure 5.6	Sample item from the recognition test	147
Figure 5.7	Overview of the learning conditions implemented at three institutions	151
Figure 6.1	Examples of structural regularities in the experimental song	183

List of Appendices

Appendix A: Materials for Study 1 (the teacher survey)

Appendix A.1: Complete questionnaire (paper-based version) Appendix A.2: Ethics approval for the teacher survey

Appendix B: Materials for Study 2 (the pedagogical song corpus)

Appendix B.1: A list of ESL textbooks used as sources for the compilation of the song corpus

Appendix B.2: A list of ESOL-teacher websites used as sources for the compilation of the song corpus

Appendix B.3: A list of lexical vocables found in the song corpus

Appendix B.4: A list of songs included in the pedagogical song corpus

Appendix C: Materials for Study 3 (the intervention study)

Appendix C.1: Ethics approval for the quasi-experimental intervention study Appendix C.2: Information sheet for participants in the intervention study Appendix C.3: Consent form for students participating in the intervention study Appendix C.3: Gap-fill activity (song version) Appendix C.4: Gap-fill activity (prose version) Appendix C.6: Free recall test Appendix C.7: Cued recall test Appendix C.8: Recognition test (song version)

Chapter 1 Introduction

1.1 Introduction: Why study the use of pop songs in the language classroom? Many language teachers express enthusiasm regarding the use of pop songs in the classroom as a tool to foster language acquisition, even for adolescent and adult language learners. A keen and unceasing interest in the matter is exemplified by the continuing publication of pedagogical articles on the use and usefulness of pop songs for second language (L2) teaching and learning over the years (e.g. Stocker, 1923; Richards, 1969; Gelman, 1973; Jolly, 1975; Abrate, 1983; Baechtold & Algier, 1986; Murphey, 1987; Dethier, 1991; Arleo, 2000; Foncesa Mora, 2000; Kao & Oxford, 2014), many with confident titles such as "Justified and ancient: Pop music in EFL classrooms" (Domoney & Harris, 1993), "Connecting the powers of music to the learning of language" (Adkins, 1997) and "Music works: Music for adult English language learners" (Lems, 2005). Language instructors recommend songs and activities on the internet on teacher websites and blogs, provide entire lesson plans and describe their experiences regarding the use of music for teaching purposes (see Appendix B for a list of ESL-websites). On these websites, teachers call songs a "great language package that bundles culture, vocabulary, listening and a host of other language skills in just a few rhymes" (www.eslgalaxy.com/music.htm, retrieved on 23/05/2014) and advise others to make use of the "wonderful benefits of using songs to teach ESL learners"

(www.teachingenglishgames.com/Articles/Using_ESL_Songs.htm, retrieved on 23/05/2014). They list a variety of arguments for the use of pop songs, for example their omnipresence in everyday life, the chunking of vocabulary and grammar in the lyrics, the presentation of meaningful language in context, songs' power to motivate and build confidence in students, and songs' potential as a memory aid. In addition, teachers propose a number of teaching goals such as fluency, vocabulary, grammar, syntax, listening comprehension, and classroom management, and a myriad of activities from the classic gap-fill exercise to the production of a music video.

In contrast to such enthusiasm stands the relative scarcity of songs in official curricula and published language teaching materials (Edwards, 1997; Huy Lê, 1999; Pérez Aldeguer & Leganés Lavall, 2012; Walklett, 2014). As a result, teachers' choices regarding the use of songs in the classroom are limited, regardless of their personal convictions. Reeve and Williamson (1987) accurately described the place of songs in many language classrooms as "a carrot – a reward for tired teachers and students on Friday of a long week" (p. 33). Another reason for such a reserve to exploit songs more fully as a teaching resource might be the overall dearth of empirical research exploring potential benefits of music (or a possible lack thereof) for second language acquisition (Conrad, 1991; Kanel, 1997; Lems, 2005). As Conrad (1991) argued: "Considering the paucity of empirical and theoretical study, it seems unwarranted to turn lyrics into exercise material solely on the basis of students' enthusiasm for the genre" (p. 22). To this day, only few empirical studies have explored aspects of song use in language teaching. In particular, only a very small number of experimental and quasi-experimental studies have explored the effect of songs on the retention of words, phrases and longer text (Hahn, 1972; Smith Salcedo, 2002; Salas, 2006a; Milton, 2008; Ludke, 2010), frequently with a particular focus on young children (Hahn, 1972; Medina, 1990).

The value of songs for language teaching and learning is, thus, viewed somewhat controversially. However, ultimately we know very little about how songs are actually used in the language classroom. While teachers' recommendations, lesson plans and anecdotal accounts are a valuable source of information and insight, they cannot replace a more thorough empirical investigation of teachers' beliefs and teaching practices. Language instructors are at the centre of pedagogical decision-making and the implementation of methodological approaches in the L2-classroom, and it is essential to get a better understanding of their views and experiences. Until now, however, only few studies have explored the use of songs in language teaching from the point of view of the teachers (Edwards, 1997; Bjorklund, 2002; Pérez Aldeguer & Leganés Lavall, 2012). Unfortunately, while providing some interesting first insight, the few existing studies have been limited in scope and in the number of participants. In addition, they did not explore the details of actual classroom practices. A focus on teachers, their views, thoughts and beliefs but also their teaching practices locates this study in the vicinity of

teacher cognition research, that is, research into "what teachers think, know, and believe and the relationship of these mental constructs to what teachers do in the language classroom" (Borg, 2003, p. 81).

Apart from a limited understanding of the teaching practices around songs, we also know little about the musical material used in song-based language lessons. It currently remains unclear what kinds of popular songs teachers typically select for in-class use. In particular, no study has investigated the lexical load of pop songs used in a classroom setting. Only Murphey (1989, 1990a, 1992) conducted a seminal analysis of lexical characteristics and discourse features of pop songs in the context of their potential pedagogical use in the language classroom. However, Murphey's corpus consisted of general pop songs from a top 100 chart which may or may not be used in the language classroom. An understanding of the lexical load of authentic teaching material is critical. As Webb and Rodgers (2009b) highlighted, learners are not likely to engage with material that they cannot understand. A number of corpus studies have explored the lexical load and learning potential of various text genres used in L2-teaching, such as novels (Nation, 2006), TV programmes (Webb & Rodgers, 2009b) and movies (Webb & Rodgers, 2009a). However, as Kreyer and Mukherjee (2007) noted, with the exception of Murphey's corpus study, "pop song lyrics are virtually absent from corpus-linguistic research" (p. 31).

1.2 Why study the learning of vocabulary through songs?

The present study has a particular focus on the use of songs for lexical learning. In this thesis, lexical learning refers to the acquisition, consolidation and deepening of vocabulary knowledge, including the knowledge of multi-word units. While teachers seem to use songs for a great variety of pedagogical purposes, a recurrent claim is that songs are useful to teach and learn vocabulary (Claerr & Gargan, 1984; Aquil, 2012; Kerekes, 2015), including longer phrases and expressions (e.g. Jolly, 1975; Baechtold & Algier, 1986; Griffee, 1988; Riddiford, 1998; Abbott, 2002; Esa, 2008). A review of the pedagogical literature reveals that songs are frequently attributed a mnemonic effect, which is understood to benefit the memorization of words (e.g. Murphey, 1990a; Foncesa Mora, 2000; Lake, 2002). In addition, it is assumed that using songs can benefit

the memorization of phrases and formulaic sequences (Stansell, 2005; Engh, 2013; Kerekes, 2015) and aid the acquisition of words in context. As Arleo (2000) put it: "The mnemonic value of songs, is, of course, one of the best reasons for using them in the classroom" (p. 10). While a number of experiments have provided evidence that a melody can support the retention of word lists as well as longer connected text in one's native language, only very few studies have investigated the effect of a musical medium on the retention of vocabulary and verbatim text in a foreign language (Hahn, 1972; Medina, 1990; Smith Salcedo, 2002; Salas, 2006b; Ludke, 2010). In addition, studies in second language research have largely focused on the effect of songs on vocabulary learning within a highly controlled laboratory setting or within an "idealized" classroom setting involving an unrealistic investment of time and effort and the use of uncommon teaching methods. In fact, to this point no study has investigated whether songs as they are frequently used in language teaching, rather than as they potentially should be used, have a beneficial effect on second language acquisition.

1.3 Why is vocabulary learning relevant and what does it entail?

In the past, the teaching of grammatical rules often took precedence in foreign language education. In recent years, however, it has been found that lexical knowledge is at least of similar importance, if not more important for language proficiency than grammatical knowledge. Lexical errors are more likely to disrupt communication than grammatical errors (Lewis, 2002). Rich lexical knowledge can, in fact, be considered a prerequisite to learning and applying grammatical rules (Meara, 1995; Barcroft, 2007). Vocabulary size is associated with writing quality, reading and listening comprehension (Laufer & Shmueli, 1997; Stæhr, 2008) and speaking fluency (Hilton, 2008). However, it has also been recognized that learning a sufficient number of words for adequate communication in a second language can be a daunting task. On the one hand, the challenge lies in in the sheer number of words second language learners need to acquire. Nation (2006) found that knowledge of 6,000 to 7,000 word families is necessary for adequate comprehension of spoken discourse in English, and a vocabulary size of 8,000 to 9,000 word families is required to deal with a variety of written texts. The teaching of vocabulary, thus, needs to be an integral part of L2 instruction. Nation (2013) argued that particularly highfrequency words "are so important that anything that teachers and learners can do to

make sure they are learned is worth doing" (p. 16). As we have seen above, some teachers are convinced that one way to make vocabulary more memorable is to teach them through songs.

On the other hand, the challenge of vocabulary learning lies in the "incremental nature of word learning" (Schmitt, 2008, p. 343) and the variety of aspects that are involved in word knowledge. While knowing a word is frequently equated to having formed a first form-meaning connection, full receptive and productive knowledge in both spoken and written form comprises various and often gradual steps including a word's grammatical functions, its collocations and its constraints of use (Nation, 2013). Many words can thus be considered partially learnt, that is, learners have mastered some aspects of a word but not others. The full acquisition of a word requires repeated encounters in order to consolidate existing and acquire further elements of word knowledge. Schmitt (2008) emphasised that "recycling has to be consciously built into vocabulary learning programmes, and teachers must guard against presenting lexical items once and then forgetting about them, or else their students will likely do the same" (p. 343). In addition, many aspects of word knowledge, such as its grammatical and pragmatic functions and its collocations and associations, cannot be acquired by studying words in isolation but require a contextualised presentation of target vocabulary in both written and spoken form.

1.4 Why is it relevant to remember verbatim text and learn multi-word units?

So far, I have described vocabulary learning predominantly as the acquisition of individual words. However, it is now widely understood that our mental lexicon consists to a large extent of longer multi-word items or lexical chunks that are processed, stored and produced as holistic units. These lexical chunks can be fully fixed, in which case they do not require (nor allow) any grammatical analysis or construction processes, or they can be semi-fixed and still require the insertion of specific grammatical and semantic detail (cf. Wray, 2002). Multi-word units comprise social formulae (*Nice to meet you!*), collocations (*broken heart*), phrasal verbs (*to dream of*), idioms (*leave no stone unturned*) and other types of formulaic language (cf. Boers & Lindstromberg, 2009). Particularly the so-called Lexical Approach (Lewis, 1993) proposes that a large percentage of our spoken

and written discourse consists of prefabricated chunks rather than original strings of arbitrarily assembled words. Such a view solves the "puzzle of nativelike fluency" (Pawley & Syder, 1983), which refers to native speakers' ability to produce fluent discourse at a rate that would exceed our cognitive capacities if all discourse did indeed consist of entirely original strings of individual words. In addition, it addresses the "puzzle of nativelike selection" (Pawley & Syder, 1983). Nativelike selection refers to the fact that native speakers do not make full use of all word combinations possible in a language that is ruled by syntactic restrictions alone. Instead, both spoken and written discourse consist to a large extent of prefabricated phrases and commonly used expressions. This has important implications for L2 learners, who often wonder whether they can express an idea in a certain way only to hear "you could, but you wouldn't" (Lewis, 2002, p. 17). Research has demonstrated that non-native speakers are capable of exploiting the processing advantage of formulaic language (Conklin & Schmitt, 2008) and that L2learners are perceived to be more fluent if they make use of multi-word units in their speech (Boers, Eyckmans, Kappel, Stengers, & Demecheleer, 2006). Consequently, L2learners need to encounter and re-encounter words in context rather than in isolation. In addition, it appears that a large part of acquiring lexical chunks is based on memorization. As Boers and Lindstromberg (2005) highlighted, "given the enormous size of the natural language's resource bank of multi-word expressions, memorizing even a limited number of these must be a tremendous task. It follows that learners would benefit from effective mnemonic techniques that are adapted to the task at hand" (p. 234). In other words, L2-learners would profit from teaching methods that support the retention of longer chunks of verbatim text. As previously mentioned, songs are assumed to benefit the memorization of longer phrases and formulaic language.

1.5 The aim and significance of the present study

In response to a dearth of empirical research, the present study was designed to examine three aspects of song use in in the L2-classroom: the teaching practice, the material used and the learning outcome. This thesis firstly explores song use from the teachers' perspective and investigates teachers' cognitions as well as their teaching practices involving songs. Secondly, it assesses the vocabulary load and other typical characteristics of songs used by language instructors for pedagogical purposes. Thirdly,

the present study examines whether songs as they are typically used can benefit verbatim text retention. Overall it attempts to answer the following question: Do songs as they are currently used in language classrooms benefit lexical learning? The study is significant in several ways: It is concerned with the use of songs for pedagogical purposes and is intended to lay some long-needed groundwork regarding a teaching technique that is frequently embraced by teachers but unfortunately scarcely researched. The present study is firmly grounded in classroom reality and is, thus, deliberately located at the interface of teaching practice and research. It thus adds a new perspective to the small body of existing research, as other studies have used either an experimental laboratory design or a classroom-based design that put less emphasis on ecological validity.

At the same time the study is intended to provide language instructors with much needed insight into the value of songs for vocabulary acquisition. A review of the pedagogical literature shows that teachers use songs in an attempt to facilitate language learning and to render the process less daunting, more meaningful and "more fun" for their learners. Especially the acquisition of vocabulary can seem overwhelming, given the sheer number of words and phrasal lexical items necessary for adequate communication. Schmitt (2008) referred to vocabulary learning as "one of the greatest hurdles facing learners in acquiring English" (p. 332). This study tries to support practitioners in their laudable efforts to facilitate vocabulary learning by assessing a particular technique, the use of popular songs, and by providing further insight into its benefits for lexical learning.

1.6 Organisation of the thesis

This thesis consists of three individual studies, a teacher questionnaire, a song-corpus analysis and an intervention study, which inform and build on each other. The three studies respond to different research questions and apply three distinct methodological approaches. Chapter 2 discusses the existing literature in one collective review, which is divided into three clearly distinct sub-sections. The three individual studies are presented and discussed separately, including a separate set of research questions, a separate methodology, data analysis, discussion of results and limitations section. Figure 1 depicts how the three studies inform each other.

Figure 1.1: The interconnection of the three studies investigating three different aspects of song use (teaching practices, material, learning outcome)



As Figure 1.1 shows, Chapter 3 presents a teacher questionnaire, which explored teachers' cognitions and teaching practices regarding the use of songs in the language classroom. Chapter 4 focuses on the compilation and analysis of a pedagogical song corpus to explore the lexical profile and other typical characteristics of songs used in the English classroom. This corpus study is, to a large extent, informed by song recommendations and information provided by the questionnaire informants. Chapter 5 discusses a quasi-experimental intervention study, which investigates the effect of songs on English learners' retention of verbatim text when used in the context of a typical language lesson. Ecological validity of the intervention is of high priority, and the design of the treatment lesson is informed both by the findings of the teacher questionnaire and the results of the corpus study. Chapter 6 finally integrates the results of all three studies and discusses the main findings of the research project as a whole, including pedagogical and methodological implications, the limitations of the overall project and suggestions of future research.

Chapter 2 Literature review

2.0 Introduction

The present thesis addresses one overall research question: Can songs as they are currently used in the language classroom benefit lexical learning? In order to answer this question, three different aspects of song use in second language teaching, the teacher, the material and the learning outcome, are explored by means of three separate studies utilizing different methodological approaches. In Study 1 an international teacher questionnaire explored why (not) and how language teachers currently use songs in class. Study 2 comprises the compilation and analysis of a pedagogical song corpus. It investigated the vocabulary load of teacher-selected popular songs regarding their lexical demand as well as the vocabulary learning opportunities they afford. In addition, the vocabulary load of song lyrics was compared to that of other authentic text genres such as novels, films and spoken discourse. Study 3 examined a potential mnemonic effect of songs and also poems compared to prose texts by means of a quasi-experimental intervention study.

Due to the distinct topics and methodological approaches of the three studies, the literature review is sub-divided into three parts. The first part reviews the small body of existing survey studies using a questionnaire to explore current teaching practices involving songs. The second part explores research into the lexical characteristics of songs used for language teaching. In addition, this section discusses studies focused on the vocabulary load of other authentic English text genres. In particular, it details the methodology used in these studies to measure lexical demand, as the present song corpus study follows a similar approach. The third part of this chapter reviews literature relevant to the quasi-experimental intervention study. Firstly, it gives an overview of memory theories relevant to an understanding of a possible mnemonic effect of songs. Secondly, it reviews pertinent research in the field of cognitive psychology studying the effect of songs and poems on verbatim text retention by native speakers. Finally, it explores the small body of research focused on second language learners' lexical learning through songs.

2.1 Study 1: Surveying teacher cognitions and classroom practices

As Chapter 1 has shown, language practitioners are often enthusiastic about using songs. Such a positive attitude contrasts with the relative scarcity of songs in formalised teaching materials (Edwards, 1997; Huy Lê, 1999; Pérez Aldeguer & Leganés Lavall, 2012; Engh, 2013; Walklett, 2014). In addition, only few empirical studies have explored the use of songs in second language teaching (see Conrad, 1991; Kanel, 1997; Lems, 2005). Due to the dearth of prescribed principles and recommended teaching techniques regarding the use of songs in second language teaching, pedagogical choices and their implementation depend mainly on the course instructor. It, therefore, seems warranted to go to the source and explore song use from the point of view of teachers. Consequently, the present survey study was designed to explore if, why (not) and how teachers currently utilise songs in the language classroom, with a particular focus on the teaching of vocabulary. In the following review, I discuss the small number of survey studies which have explored the use and usefulness of songs in the language classroom from the perspective of the teacher. Following the review, I highlight the research gap that the present study intended to fill.

2.1.1 Exploring teachers' beliefs and classroom practices

To this day, only few studies have researched the actual use of songs in the language classroom in a systematic way. Particularly three survey studies have focused on the teacher perspective in an attempt to provide descriptive information about the use of music and songs in foreign and second language teaching and the teachers' reasoning behind their instructional choices. Only one of these studies focused on teachers working with adult learners. Bjorklund (2002) conducted a teacher survey for her Master's thesis in the USA. She focused on the use of music to teach oral skills, that is, pronunciation, stress and intonation, to adult ESL learners. A paper-based and a digital version of a questionnaire were distributed among ESL instructors at the University of Wyoming. Additionally, a digital version was distributed to a listserv for TESOL graduates. In total, 30 teachers responded to the survey. Bjorklund combined a quantitative and a qualitative approach and included questionnaire items with prefabricated as well as open-ended answer options. Overall, Bjorklund found that the majority of her informants (87%) used music to teach certain oral skills. The most popular type of music were pop

songs, utilised by 70% of music-users, followed by clapping activities (65%) and to a much lesser extent rhythm games and jazz chants (both 30%). The majority of respondents (78%) reported observing improvements in their students' oral skills after engaging with musical activities. In addition, most music-users (88%) reported a positive response by their students, with only two respondents describing negative learner reactions.

While Bjorklund's research is the only systematic study assessing the actual use of music and songs with adult second language learners, it is unfortunately fairly narrow in scope. Bjorklund only focused on the question whether teachers used music, and if yes, what types. She did not explore the informants' classroom practices in greater detail. In addition, the study did not investigate the teachers' beliefs and knowledge regarding the use of music for pedagogical purposes. Overall, Bjorklund's findings are of limited generalizability to the wider population of language teachers because of the small number of informants and also due to the respondents being representative of only a very specific group, that is, ESL instructors working mostly in tertiary education in a North-Western state of the USA. Finally, Bjorklund's study design showed that the fundamental research concepts and their equivalent technical terms need to be carefully defined: As the author acknowledged, her definition of the term "song" remained vague. She frequently used the term to refer to music in general rather than using the word in the widely accepted way of referring to a combination of music and lyrics. Bjorklund acknowledged that this might have confused informants.

Two studies, Edwards (1997) and Pérez Aldeguer and Leganés Lavall (2012), investigated the use of songs by teachers working with children. Edwards (1997) conducted a survey for her Master's thesis in one school district in California, USA, in order to describe possible ways of using music in teaching English as a second language (ESL). In a paperbased questionnaire, Edwards asked 33 teachers (28 female, 5 male) of young ESL learners (kindergarten to grade 5), if and for what purposes they used music in the classroom. In addition, she inquired what prevented them from using music. Despite its focus on primary school teachers only, Edwards' study offered greater insight into teachers' cognitions and classroom practices related to songs. Edwards found that over two thirds of her informants used music to teach English and that their three primary purposes were the teaching of vocabulary, the lowering of the affective filter (Krashen &

Terrell, 1983) and the improvement of cultural awareness and appreciation, all selected by 88% of informants. Other purposes proposed by Edwards in a checklist, such as reading comprehension, pronunciation, syntax and grammar, were selected to a much lesser extent by 44% or fewer of the informants. Edwards' respondents understood a lack of money for supplies (88%) as well as a lack of training in the use of songs (76%) to be the main obstacles to using music in the classroom, while one third of informants pointed out a lack of time to include music in the daily schedule. When asked how to encourage a greater use of music, informants responded in an open-answer format, suggesting the provision of more funds, of equipment and material and the provision of further training in the form of workshops and additional opportunities to share ideas. One limitation of Edwards' study is the small number of very specific informants, that is, teachers working in a particular setting with a particular group of learners. Her findings, while interesting, cannot easily be generalised to a wider population of teachers and classrooms.

Pérez Aldeguer and Leganés Lavall (2012) took a greater interest in teacher cognitions than previous survey studies. They sent a questionnaire with 15 closed-answer items to 20 primary school instructors of English as a foreign language (EFL) at 20 schools in the Comunidad Valenciana in Spain. They investigated (1) the informants' awareness of various benefits of music in EFL instruction, (2) their opinions regarding the potential of music in EFL teaching, (3) their use of music in their classes, and (4) their interest in didactic material involving music. Pérez Aldeguer and Leganés Lavall found that half of their informants reported being aware of pedagogical possibilities and a majority (85%) was aware of the motivating effect of music. However, only 30% of the respondents reported using music with all their classes. When asked about their knowledge of materials and teaching techniques, only 25% of informants confirmed that they knew how to use musical activities inside the classroom, and only 30% reported knowing a good selection of songs and rhymes in English. Only half of the informants reported that they enjoyed using musical activities in their English classes. The authors concluded that the majority of informants showed a great interest in musical teaching materials and had a positive attitude towards music in the EFL classroom but did, in fact, scarcely make use of it. Like Edwards (1997), the authors argued that a frequent obstacle to the use of music was the teachers' lack of knowledge of how to use songs to facilitate learning and

how to integrate musical activities into the curriculum. Consequently, Pérez Aldeguer and Leganés Lavall called for the development of more teaching material involving music that clearly states didactic goals and teaching techniques.

Pérez Aldeguer and Leganés Lavall's research on the use of music by EFL teachers in Spain provides a welcome addition to the existing survey studies that so far focused on teachers in the USA alone. One limitation, however, is the order as well as the wording of some of the questionnaire items which may have biased the informants and resulted in a *halo effect* (Thorndike, 1920). In other words, the respondents might have, to a certain extent, provided answers that they understood to be preferred by the researchers. Examples of questions that might have influenced the informants' response behaviour are "I am informed about the benefits provided by the use of musical activities ..." and "I am aware of the motivation that the use of musical activities can instil in my students". In addition, some of the questionnaire items could be considered double-barrelled or at least difficult to interpret.

As the literature review shows, only few studies have so far concerned themselves with the teachers' views and their actual classroom practices regarding the use of songs in the foreign language classroom. The three reviewed studies have provided some indication that language teachers, working with both children and adults, have a positive attitude towards songs and music as teaching tools. However, regarding the actual use of music, the three studies show a diverse picture: They provided some evidence that teachers consider a lack of official prepared materials, a lack of funding and a lack of specific training to be major obstacles to the use of music in class. However, while these obstacles seemed to prevent a majority of Pérez Aldeguer and Leganés Lavall's EFL teachers in Spain from implementing musical activities, these problems did not seem to stop Edwards' informants in the USA from integrating songs into their lessons.

When it comes to the details of classroom practices, the existing research provides little insight: Only Edwards explored the teachers' instructional goals. Relevant to the present study, Edwards found the teaching of vocabulary to be a frequent pedagogical purpose among her informants. However, overall teachers' reasoning behind the use or non-use of songs remained largely unexplored. Furthermore, none of the reviewed surveys

investigated the impact of internal and external factors on the teachers' decision-making. While all of the studies discussed above obtained demographic and personal information about their informants such as gender, teaching experience and musical preferences, they made little use of this data. In general, it is difficult to provide a comprehensive picture of the matter at hand, as all three studies focused on different aspects: They targeted teachers working with young children (Pérez Aldeguer & Leganés Lavall; Edwards) or adults (Bjorklund); they investigated the use of music for various purposes (Edwards; Pérez Aldeguer & Leganés Lavall) or for only one particular purpose, that is, the teaching of oral skills (Bjorklund); they focused on a small, specific cohort in the USA (Edwards; Bjorklund) or in Spain (Pérez Aldeguer & Leganés Lavall).

2.1.2 Implications of previous research for the current survey study

All in all, existing research still falls short of providing a comprehensive and in-depth understanding of what teachers believe and do when it comes to songs as language teaching tools. It is currently unclear whether the enthusiasm conveyed in many pedagogical articles (see Chapter 1), on websites and in personal conversation with language instructors is indeed widespread among teachers and does in fact result in the use of songs in language classrooms. The present study, therefore, systematically investigated the views and practices of teachers in relation to songs by means of an international teacher questionnaire. The questionnaire was targeted at a wide range of teachers working in various cultural and institutional settings. Woods (1996) made the criticism that "information about teaching has generally been normative rather than descriptive" (p. 21). However, as songs are not usually part of formalized teaching, it is of even greater importance to investigate their use in an exploratory and descriptive manner. The questionnaire, thus, was intended to provide a clear picture of how exactly songs are being used in the classroom. In addition, teacher cognition research has found that teachers' views and classroom practices are influenced by a number of factors, including institutional and curricular constraints, their professional coursework, their ongoing teaching experiences but also experiences made when they themselves were students (Johnson, 1994; Borg, 2003; Ellis, 2006). Consequently, the questionnaire in the present thesis explored internal and external factors that may influence teachers'

instructional choices related to songs. The present survey study including its exact research questions, its methodology and its results is presented in Chapter 3.

2.2 Study 2: The lexical profile of songs used in the language classroom

One important aspect of a musical teaching approach is, of course, the song used as material. When selecting a song, teachers must bear in mind a myriad of factors, including the suitability of both its content and its language for their students. In particular, teachers need to consider the vocabulary used in the lyrics and whether it is appropriate for the target learners (Richards, 1969; Jolly, 1975; Murphey, 1990a; Purcell, 1992; Smith Salcedo, 2002; Aquil, 2012). When discussing the suitability of vocabulary, one relevant question is: How many words do learners need to know to gain adequate comprehension of authentic English songs used as teaching material? In addition, when considering songs as a means to teach vocabulary, it is relevant to know the lexical profile of songs and how it can best serve lexical learning. Finally, songs are one text genre among many for teachers to choose from. It, therefore, needs to be considered how the lexical profile and particularly the vocabulary load of song lyrics and their vocabulary learning potential compare to that of other authentic language materials such as novels, TV programmes or movies. Consequently, the present study followed the methodology of existing research such as Nation (2006), Webb and Rodgers (2009a) and Webb and Rodgers (2009b), which used a corpus-based approach to gauge the vocabulary load of various types of written and spoken discourse. The use of the same methodological approach allowed for a direct comparison of song lyrics with other text genres. In the following literature review, I first discuss existing corpus-based research exploring the lexical profile of songs. I then present existing studies examining the vocabulary load of other types of written and spoken discourse and their methodological implications for the study at hand.

2.2.1 Existing research on the lexical profile of songs in language teaching

To this day, only one corpus study, Murphey (1990a; also 1989, 1992), has investigated pop songs from a language-pedagogical point of view. In his seminal research project, Murphey compiled a corpus of 13,161 tokens from 50 English songs taken from the *Music* & *Media* Hot 100 Chart in the second week of September 1987. This chart ranked the

best-selling songs in 18 European countries during that week including several hits by Madonna, the Pet Shop Boys and Whitney Houston. Murphey conducted a lexical analysis as well as a content analysis of the lyrics. The content analysis revealed that a majority of the lyrics used first- and second-person personal pronouns and possessive determiners. Typically, these personal referents did not specify a particular person, remained vague about the referents' gender and even often displayed a certain fluctuation over who was the subject and who the addressee. Furthermore, Murphey conducted an analysis of the times and places referred to in the lyrics and discovered that most songs did not mention a specific time or location. These findings and the fact that the songs comprised a high number of imperatives and questions led Murphey to conclude that pop song lyrics are similar to spoken situational discourse.

Murphey (1989) argued that vagueness regarding persons, time and location prevalent in pop songs allowed for the listener to "complete the message, or make sense of the song, through using the persons, times and places from their own physical and metaphysical situation" (p. 185). In effect, he argued that listeners can appropriate the lyrics as their own. Consequently, pop lyrics can be seen to be highly interesting and emotionally appealing particularly but not only to adolescent listeners. In this context, Murphey also raised the question whether pop songs should be considered written or spoken discourse. This question is discussed in greater detail below.

Apart from a content analysis, Murphey (1990a) also conducted a lexical analysis of his song corpus, which is of greater interest to the present study. Murphey showed that the average song in the corpus contained 263 words. In addition, he found that the song corpus displayed a type-token ratio (TTR) of .087. The average song had a type-token ratio of .29, implying fairly high repetition of a relatively limited number of words. Murphey compared this result to an excerpt of similar length (394 words) from an English textbook for 1st-year-students and found that the textbook passage had a clearly higher TTR of .37. Murphey also calculated the frequency of each word in the corpus and found that only ten words, *you, I, me, my, the, to, a, and, gonna, love* (as both noun and verb) accounted for 25% of the total number of tokens in the corpus. Finally, he analysed the lexical and syntactical complexity of the song lyrics and found that they contained comparatively short sentences with an average sentence comprising seven words. He

also calculated that 82% of tokens in the corpus were mono-syllabic words. Based on the results for sentence- and word-length, Murphey applied Flesch's (1974) readability formula to the song corpus and found that lyrics can be considered "very easy" (on a 7-point-scale from very difficult to very easy), comparable to texts requiring a reading level of a child after five years of schooling. Finally, Murphey also calculated the speech rate in songs and found that it was about half the speed of normal spoken discourse. However, according to Murphey, such a low speech rate did not imply a slower vocalization of words but instead was mostly due to a high number of pauses. From his lexical analysis, Murphey concluded that pop songs as a genre are short, lexically simple and highly repetitive. His findings support Murphey and Alber's (1985) previously stated hypothesis – based on a small-scale study of one song only – that pop lyrics are simple, non-threatening, highly affective and, in essence, a "motherese for adolescence" (p. 794), and as a result might lend themselves well for language learning.

Murphey's findings provided a good indication of the results that could be expected from the present song corpus study. However, findings in the current study were also expected to differ from Murphey's for two reasons. Firstly, the present corpus is considerably larger comprising a significantly higher number of tokens from a greater sample of songs. Secondly, the current corpus comprises songs selected by teachers and material designers for in-class use rather than a selection of popular hits from a randomly selected chart. The lexical profile of the present corpus, thus, not only reflects the typical characteristics of pop songs as a genre, but is also influenced by the languagepedagogical decisions made by educators. In addition, Murphey's lexical analysis displays a small number of limitations that the current study set out to address.

Firstly, Murphey's claims about the lexical simplicity of pop song lyrics are partially based on his finding that only few words are highly frequent and make up a large part of the corpus. However, research has since shown that high-frequency words such as *the*, *and*, *to*, and *you* account for large numbers of tokens in any form of discourse and are not necessarily an indication of a particularly simplistic type of language. As described by Zipf's law (Zipf, 1949), words decrease rapidly in frequency. Nation (2006) found that the most-frequent 1,000 word families and proper nouns in English account for 78% to 81% of written text and about 85% of spoken text. Words from the 10th to 14th frequency

levels together, on the other hand, only account for about 0.5% to 1% of tokens. It, thus, seems that Murphey's claim of song's lexical simplicity still requires further empirical support.

Secondly, Murphey's conclusion regarding the high lexical repetition in song lyrics is based on the analysis of the type-token ratio. The type-token ratio as a measure of lexical diversity is highly dependent on the length of a text or the size of a corpus. As a result, it is limited in its informative value and can only be interpreted in direct comparison with a text of similar length. The present study made use of the so-called standardised typetoken ratio (STTR) found in WordSmith Tools® (Scott, 2008) as a measure of lexical diversity. The STTR is computed by calculating the average of each TTR for every consecutive chunk of 1,000 tokens in a text (for further details see http://www.lexically.net/downloads/version5/HTML/index.html?type_token_ratio_proc.htm, retrieved 15/11/2014). The standardised type-token ratio allows for a comparison of texts of varying lengths.

Thirdly, Murphey had no means to gauge the vocabulary load of pop songs as a genre. The present study will make use of a methodological approach not available when Murphey conducted his song corpus analysis. This methodology uses word frequency lists to measure the overall lexical demand of a corpus. It has been used in a number of studies on the vocabulary load of various types of written and spoken discourse and is explained in greater detail below. As an additional advantage, the use of a methodology applied in the study of various text genres permits a comparison of the overall lexical demand of different teaching materials. However, it also needs to be acknowledged that applying this methodology to the analysis of songs holds certain limitations. In summary, research into the lexical profile of songs is scarce and leaves several questions unanswered. However, in recent years, a growing number of studies have investigated the lexical load of various written and spoken types of English discourse from a languagepedagogical perspective. In the following section, I will discuss the methodology applied in these studies as well as their findings for different text genres.
2.2.2 Methodological considerations: Assessing the lexical demand of a text genre

An increasing number of studies within the field of second language vocabulary research are concerned with the overall lexical demand of various text genres from a languagepedagogical perspective, such as novels, newspaper articles, movies and TV programmes. They are interested in how many words second language learners need to know for adequate comprehension of these types of written and spoken discourse and for incidental vocabulary learning to occur when engaging with these text types. In order to assess the lexical demand of a text genre, one has to address two questions. Firstly, one has to determine a reliable average percentage of how many words in a text need to be known for a reader to gain adequate comprehension of its content. Secondly, we need to have an understanding of how many words a learner is required to know to reach that threshold.

2.2.2.1 How many words of written and spoken texts need to be known for adequate comprehension?

Within vocabulary research, the issue of lexical demand is frequently addressed by asking how many words in a text need to be understood for adequate or reasonable comprehension (Nation, 2006; Stæhr, 2008; Webb & Rodgers, 2009b; Laufer, 2013) and for incidental vocabulary learning to occur (Webb & Rodgers, 2009a). Coverage of around 95% to 98% of words in a target text has been suggested (Laufer, 1989; Hu & Nation, 2000; Schmitt, Jiang, & Grabe, 2011). In this context, the term coverage refers to the percentage of known words in the text. 98% coverage is widely accepted to be the optimal threshold (Laufer, 2013) for adequate comprehension of unsimplified written texts. That is, with knowledge of 98% of running words or tokens in a target text learners are likely to reach acceptable understanding of the content. However, while it has been repeatedly demonstrated that 98% coverage is required for optimal reading comprehension (Hu & Nation, 2000; Laufer & Ravenhorst-Kalovski, 2010), the same threshold cannot simply be applied to listening comprehension. Listening differs from reading in many ways, most obviously in the temporary nature of aural texts and the challenge of parallel reception and decoding. Aural texts in contrast to written input do not provide opportunities for perusal and repetition (Lund, 1991). However, spoken discourse also affords extra-linguistic support to understanding. It frequently provides

non-verbal clues, such as gestures, facial expressions and lip movements, which aid listening comprehension and make up for deficient lexical knowledge (van Zeeland & Schmitt, 2013). Given the additional non-verbal cues present in many forms of spoken discourse, van Zeeland and Schmitt (2013) argued that coverage necessary to comprehend written and spoken texts might differ. Based on an experimental study of listening comprehension in native and non-native speakers, they proposed 95% as an appropriate coverage target for listening comprehension of informal spoken narratives. Bonk (2000) investigated EFL-learners' comprehension of four audio-recordings of varying levels of lexical difficulty and concluded that coverage of less than 95% of tokens might still result in adequate comprehension if listeners made use of effective coping strategies. In contrast, Stæhr (2008) found that 98% coverage seemed to be a reasonable threshold for adequate listening comprehension. And Nation (2006) cautioned that "we need to consider that text coverage greater than 98% may be needed to cope effectively with the transitory nature of spoken language" (p. 79). As this short overview reveals, there is currently no consensus regarding the optimal lexical coverage of aural texts. Song lyrics, however, can be considered a fairly particular text genre, displaying characteristics of both written and spoken discourse (Kreyer & Mukherjee, 2007). The textual characteristics of and the appropriate coverage for song lyrics are further explored below.

2.2.2.2. How many words do learners need to know to reach optimal lexical coverage?

The next issue that must be addressed when assessing the lexical demand of a text or text genre is the question of how many words learners need to know to reach the threshold required for adequate comprehension. This is often done by assessing the coverage of a text provided by word frequency lists. In this context, *coverage* refers to the percentage of words accounted for by such word lists (Nation, 2004). Frequency lists sort words according to their frequency in general language use, from most to least frequent. Nation's (2004, 2006) BNC frequency lists, for example, rank English words according to their frequency, range and dispersion in the British National Corpus (BNC). Using such frequency lists to assess the vocabulary knowledge required to understand various text genres is based on the assumption that language learners acquire common words earlier than less common vocabulary. Research has shown that this is indeed the

case (Nation, 2006). Nation (2006) explained that "high-frequency and wide-range words are generally learned before lower-frequency and narrow-range words" (p. 63).

It needs to be noted that the BNC wordlists (Nation, 2004, 2006) used in the majority of studies on lexical coverage consist of word families rather than individual words. That is, the lists contain headwords along with a number of family members. A vocabulary size of 3,000 word families, consequently, refers to knowledge of more than 3,000 individual words, as each word family can comprise several family members. For the BNC wordlists, a word family is defined on the basis of the level-6 classification described in Bauer and Nation (1993), which includes inflected and derived forms. The headword break, for example, subsumes the following members: breaking, breakage, breakages, breaks, broke, broken, unbreakable, outbreaks, outbreak, unbroken, breaker, breakers. The use of the word family to measure word knowledge is based on the assumption that "inflected and regularly derived forms of a known base word can also be considered as known words if the learners are familiar with the affixes" (Hirsh & Nation, 1992, p. 692). In addition, proper nouns are also often assumed to be known or have such a small learning burden as to be counted as known (Hwang & Nation, 1989; Hirsh & Nation, 1992). Socalled marginal words are also often counted as known due to their low learning burden (Nation, 2006). The term marginal words refers to exclamations, interjections and hesitation markers such as oh, err, or ah. Finally, Nation has added transparent compounds as a separate category to the word frequency lists, as they can also considered to be known by means of knowing their high-frequency parts, for example hometown, babygirl or sailboat. See Bauer and Nation (1993) for an explanation of transparent compounds and (Nation, 2014) for an application of the category in a corpus study.

2.2.3 The vocabulary size necessary to comprehend various text genres

No study to date has attempted to establish the lexical demand of pop song lyrics based on the coverage provided by word-frequency lists. In contrast, various studies have explored the lexical demand of other text genres used in English teaching. Table 2.1 shows the word knowledge required for 95% and 98% coverage of different text genres, if available. As this research project is focused on the use of authentic songs, that is, songs composed for native speakers, the literature review is limited to studies

investigating authentic texts and excludes research focused on materials specifically written for language learners such as graded readers.

Text Genre	95%	98%	Study
Songs	?	?	
Novels (adult)	4-5,000 + PN	8-9,000 + PN	Nation 2006
Novels (adolescent)		5,000	Hirsh & Nation, 1992
Newspaper articles	4,000 + PN	8-9,000 + PN	Nation, 2006
Spoken discourse	3,000 + PN	7,000 + PN	Nation 2006
Spoken discourse	3,000		Adolphs & Schmitt, 2003
Movies	3,000 + PN, MW	6,000 + PN, MW	Webb & Rodgers, 2009b
TV programmes	3,000 + PN, MW	7,000 + PN, MW	Webb & Rodgers, 2009a

Table 2.1: Lexical knowledge, indicated in word families, required for 95% and 98% coverage of different authentic text genres (PN = proper nouns, MW = marginal words)

As Table 2.1 shows, figures differ overall when it comes to written, spoken and scripted spoken text genres. In the following, I will discuss the existing research for these three categories of discourse separately.

2.2.3.1 The lexical demand of written texts

Research into lexical coverage has originated from and is still primarily focused on written texts and reading comprehension. Nation (2006) investigated the vocabulary demand of various types of written discourse. He found that vocabulary knowledge of 8,000 to 9,000 word families is needed to read novels. This is a significantly higher demand than the 5,000 word families first proposed by Hirsh and Nation (1992). One reason for this difference could be the use of different wordlists: Hirsh and Nation based their analysis on West's (1953) General Service List (GSL) and on wordlists adapted from Thorndike and Lorge's (1944) vocabulary workbook, whereas Nation (2006) used fourteen wordlists developed from the BNC. Secondly, whereas Hirsh and Nation analysed three short novels written for adolescents, Nation (2006) investigated the vocabulary demand of five novels written for adult readers. These five novels not only constituted a larger corpus from a greater variety of sources but they were also written for a more sophisticated audience. For these two reasons, Nation's corpus might have

displayed greater lexical diversity. Finally, in Nation's study there might have been a greater mismatch between the analysed texts and the wordlists used to analyse them: Nation's corpus only contained novels published before 1930, yet the wordlists used where based on the BNC, a corpus "designed to represent a wide cross-section of British English from the later part of the 20th century"

(http://www.natcorp.ox.ac.uk/corpus/index.xml, retrieved on 23/07/2014). The comparison of Hirsh and Nation's (1992) study of novels for adolescents and Nation's (2006) analysis of novels written for an adult readership highlights the importance of a precise definition of the targeted text genre. It also emphasises the necessity for a corpus to contain a comprehensive selection of text samples if it is intended to be representative of a genre. As previously discussed, the only existing pedagogical song corpus (Murphey, 1990a) contained a fairly small number of songs and comprised several songs produced by the same artists, thus possibly over-emphasizing a particular style of language.

Nation (2006) also investigated the lexical demands of various newspaper corpora and found that they required the same level of vocabulary knowledge as novels. That is, 8,000 to 9,000 word families plus proper nouns were necessary to gain adequate reading comprehension of newspaper articles. This might be surprising, as one might expect a higher vocabulary demand in newspaper articles, which can be assumed to cover a wider variety of specific topics and to use a greater number of academic terms than novels. However, there are, in fact, clear differences between the lexical profiles of the two genres: The first 2,000 word families provided coverage of 87.8% of tokens in novels but only of 83% of running words in the newspaper corpora. And whereas, proper nouns overall made up only 1.53% of words in the novels, they provided coverage of between 4.6% and 6.1% of tokens in newspaper articles. This is, as Nation (2006) argued, "not surprising, because newspapers are about people, places, and events" (p. 72). Their high percentage of proper nouns, thus, lowered their overall lexical demand. This finding highlights the necessity to take into account not only the overall vocabulary load but also the particular lexical profile of a genre. Interestingly, song lyrics seem to contrast with newspaper articles as they tend to be deliberately vague about people, places and times (Murphey, 1990a). It can thus be expected that song lyrics contain a comparatively low number of proper nouns.

2.2.3.2 The lexical demand of unscripted spoken texts

Unscripted spoken discourse, that is, spontaneous, unplanned speech not based on a script prepared prior to speaking, has been found to be less demanding in terms of vocabulary load than written text genres. Nation (2006) analysed two parts from the spoken section of the Wellington Corpus of Spoken New Zealand English (WSC) of around 100,000 words each and found that knowledge of 6,000 to 7,000 word families plus proper nouns were necessary to understand 98% of words of friendly conversations and talk-back radio or interviews. Nation conceded that the use of wordlists based on a British corpus might have led to a somewhat inflated result, as some highly frequent words in the New Zealand corpus such as Zealand, Maori or Wellington were counted as low-frequency words in the analysis. Adolphs and Schmitt (2003) investigated general spoken discourse by analysing the 5-million-word Cambridge and Nottingham corpus of discourse in English (CANCODE), which consists of speech recorded in Great Britain. They found that 3,000 word families provided almost 96% coverage of the corpus. Furthermore, they discovered that there was great variation between different types of spoken discourse. The CANCODE differentiates between five types of conversations, defined by the relationship between interlocutors: intimate, socio-cultural, professional, transactional and pedagogical. While 3,000 word families provided over 96% coverage of the transactional subcorpus and still over 95% of the professional section, it only provided around 93% coverage of the pedagogical subcorpus. Adolphs and Schmitt concluded that pedagogical discourse comprises a much higher percentage of lowfrequency formal and academic vocabulary.

It needs to be noted that Adolphs and Schmitt (2003) did not utilise Nation's BNC wordlists but instead used frequency lists based on the CANCODE itself. That is, coverage was assessed by the frequencies of the words in the corpus itself, rather than by more generalised frequency lists based on both written and spoken texts from a variety of contexts. As Nation (2006) explained, employing wordlists representative of general language use is done "to represent the vocabulary size of a typical language user. Such a user would not know only the words in a spoken corpus such as the CANCODE but would know other words as well" (p. 62).

2.2.3.3 The lexical demand of scripted spoken texts

The category of scripted spoken discourse includes lectures, speeches and language produced in movies and TV programmes. It differs from unscripted spontaneous language in that it follows a written script prepared prior to speaking. Webb and Rodgers (2009a) examined the lexical demand of British and American movies, while Webb and Rodgers (2009b) investigated the vocabulary demand of British and US-American TV programmes. The movie corpus contained 2,841,887 words from 318 movies with a total running time of 601 hours and 33 minutes, and an average running time of 113.5 minutes. The TV corpus contained the scripts of 88 television programmes with an estimated total running time of 35.1 hours, comprising 264,384 words. Both the movie and the TV corpus required knowledge of 3,000 word families plus proper nouns and marginal words to reach 95% coverage. For 98% coverage of the movie corpus, knowledge of 6,000 word families plus proper nouns and marginal words was necessary, while the TV corpus required knowledge of 7,000 word families.

Both studies found fairly great variation in their corpora: In the movie corpus, there was variation between genres, with drama and horror showing overall the lowest lexical demand, requiring 5,000 word families to reach 98% coverage. Surprisingly, animated movies displayed the highest vocabulary load with a required knowledge of 10,000 word families. Webb and Rodgers speculated that this comparatively high demand was largely due to the high use of words such as *squirrel* and *zebra*. Such words are frequent in child language but are far less common in adult conversations and consequently of low frequency in the BNC wordlists used during the analysis. Webb and Rodgers (2009b) also found great variation in the TV corpus between individual episodes of TV shows and particularly in news programmes, with necessary word knowledge ranging from 2,000 to 5,000 word families plus proper nouns and marginal words for 95% coverage. Such a great variation was, as Webb and Rodgers argued, "due, in part, to the short length of news stories" (p. 349). The shortest text in the TV corpus comprised 523 words. Song lyrics are typically even shorter. It can, therefore, be assumed that variation in lexical demand will be of even greater relevance for the song corpus analysis.

Both Webb and Rodgers (2009a) and Webb and Rodgers (2009b) are highly relevant for the present song corpus analysis, as their analysed text genres, movies and TV

programmes, share important features with songs and resemble them more closely than previously discussed types of written and spoken discourse. TV programmes, movies and song lyrics are scripted spoken language. All three genres are typically recorded and can be listened to repeatedly. In that sense, they differ from spoken discourse which Nation (2006) described as "heard and then gone" (p. 76). In addition, all three genres can be supplemented with reading material: In the case of songs, teachers frequently provide the learners with the lyrics in written form, while TV programmes and movies can be viewed with captions. Furthermore, all three genres provide additional non-verbal support for comprehension: While TV shows and movies offer strong visual support in the form of video (J. Rubin, 1994; Chapple & Curtis, 2000), songs feature music. Finally, all three genres are highly popular with learners. In other words, they are authentic genres which learners engage with voluntarily and frequently. As Webb and Rodgers (2009b) pointed out, "although we strongly advocate learning vocabulary through reading, the fact that people spend more time watching television than reading suggests that it could be an effective method of learning vocabulary" (p. 356). This claim also holds true for pop songs.

Song lyrics, then, seem to share a number of characteristics with scripted spoken discourse. However, as will be discussed in the following, they also possess features found in written and unscripted spoken discourse. As coverage thresholds for written and spoken discourse differ, it is necessary to define more closely where songs can be located within these different categories.

2.2.4 Can songs be considered written or spoken texts?

As we saw earlier, Murphey (1989, 1990a, 1992) identified a number of characteristics of lyrics indicating that pop songs should be considered situational spoken discourse. On the other hand, he found that the speech rate in songs is half that of interactional spoken discourse, which clearly distinguishes songs from other forms of spoken text. Song lyrics also display a number of characteristics more typical of written texts. Most obviously, the words are written down and usually carefully edited during the production process. And while songs can be considered transitory, they are nowadays typically recorded and can be and are usually repeated many times. Also, the provision of the lyrics in a booklet and the high popularity of lyrics websites on the internet indicate that consumers read and re-read the lyrics of songs. Kreyer and Mukherjee (2007) argued that song lyrics "sit somewhat uneasily on the boundary between writing and speech" (p. 37).

Kreyer and Mukherjee compiled the Giessen-Bonn Corpus of Popular Music (GBoP) containing 176,000 words from 442 songs on 27 top albums from the 2003 US Album Charts. They analysed the corpus regarding a number of lexical, lexico-grammatical and thematical aspects and compared them to the written and spoken sections of the International Corpus of (British) English (ICE-GB). Results indicated, among other things, that song lyrics display characteristics of both written and spoken text. The short average word-length and the high use of the personal pronouns *I* and *you* support the view of song lyrics as resembling spoken discourse. However, the comparatively high standardised type-token ratio (STTR) speaks for the categorization of lyrics as written text. While the spoken part of the ICE-GB has an STTR of 35.85, the GBoP has a higher STTR of 38.44, which indicates a lexical diversity similar to that of the written part of the ICE-GB (STTR = 38.68). In addition, Kreyer and Mukherjee observed that song lyrics in the corpus display a low use of the discourse marker you know. In interactional spoken discourse, this expression is frequently used to manage information or to keep the turn. As Kreyer and Mukherjee explained, using fillers such as you know to avoid losing the turn is "not relevant to pop song lyrics as they resemble written texts with regard to the clearly off-line text production" (p. 46). Furthermore, the authors observed that published written lyrics often display creative and unconventional ways of spelling words, as in Avril Lavigne's Sk8er boi. Such intentionally deviant spelling emphasises the relevance of the written text. Overall, Kreyer and Mukherjee identified song lyrics as "a special case of written-to-be-spoken (or, rather, written-to-be-sung) genre" (p.37). It therefore seems appropriate to analyse song lyrics using research methods that are typically applied to written texts. In addition, the identification of lyrics as belonging to both the written and the spoken genre highlights their possible usefulness for language learning, as they can be listened to as well as read, and also read-while-listening. See Brown, Waring and Donkaewbua (2008) on the benefits of reading-while-listening.

2.2.5 Implications from existing research for the present corpus study

As this literature review has shown, only one corpus study (Murphey 1989, 1990a, 1992) has conducted an analysis of the lexical properties of songs with "a language-pedagogical slant" (Kreyer & Mukherjee, 2007, p. 31). The present song corpus study intends to contribute to the existing research and to address a small number of limitations. In particular, it comprises a larger corpus with a greater sample size and uses different analytical tools. It also needs to be pointed out that the current corpus comprises songs selected by teachers and material designers to meet the needs of English learners. As a result, it differs from Murphey's (1990a) corpus of popular hits found in a European chart, and consequently its findings cannot be generalised to the genre of popular songs as a whole. Instead, the current corpus study is designed to reveal whether songs in the classroom are accessible to learners in terms of their vocabulary demand and to show what vocabulary learning opportunities songs afford. Results will indicate how lyrics should be utilised in class considering their lexical profile.

For the analysis of the song corpus, the present study made use of Nation's (2004, 2006) BNC-wordlists to assess the vocabulary load of pop song lyrics. This method made it possible to gauge the vocabulary load of song lyrics against general language use and to assess the vocabulary knowledge necessary for language learners to engage with songs. As the literature review has shown, this well-tried methodology has been utilised in the analysis of a number of written and spoken types of discourse and thus permits the comparison of various text genres regarding their lexical challenge for language learners. To this day, this methodological approach has not been applied to the analysis of song lyrics. Besides the clear benefits of this approach, there are certain caveats associated with the application of this method to the analysis of songs. Firstly, assessing the vocabulary knowledge required for adequate understanding of lyrics presupposes a recognised threshold of lexical coverage for the comprehension of songs. However, currently such a threshold has not been established. In addition, the use of BNC wordlists to measure lexical coverage is intended to assess vocabulary knowledge required for unassisted comprehension of texts and for incidental vocabulary learning. This purpose contrasts with the way songs are used in the classroom. A review of the pedagogical literature implies that songs are frequently used to intentionally teach vocabulary and

that comprehension is commonly assisted by a variety of measures, including preteaching of unknown vocabulary, explanations, repeated listening, and a variety of other form- and meaning-focused activities.

While these caveats should be kept in mind when interpreting the results of the present study, the method is nonetheless an appropriate approach to gain insight into the lexical profile of song lyrics and an improvement on methods previously applied to the analysis of song lyrics. As the literature review has shown, song lyrics resemble both written and spoken discourse. It can thus be hypothesised that lexical coverage required for adequate comprehension of lyrics is located within a range of 95 to 98 percent. In addition, lexical coverage necessary and sufficient for the comprehension of any text genre depends on a number of variables including the purpose of the activity (Stæhr, 2009) and the subcategory of a genre (Adolphs & Schmitt, 2003). As Stæhr (2009) pointed out, "the degree of lexical coverage required and the vocabulary size needed to reach this coverage will always depend on the input text and the degree of comprehension required" (p. 596). The details of the present song corpus study including its exact research questions, its methodology and its results are presented in Chapter 4.

2.3 Study 3: The benefits of songs for lexical learning

The previous two parts of the literature have explored teachers' beliefs and teaching practices involving songs and the lexical profile of songs and other text genres used as teaching material. The third part of this review takes a closer look at the learning outcome. More specifically, this section explores if and how songs might benefit lexical learning.

There is a widespread belief that songs and also poems aid the memorization of verbal information. Van Asselt (1971) reported that "older friends have shared with us German songs from their childhood or from a German class with the statement that they have retained little else but have no difficulty singing songs learned as long as fifty years ago" (p. 131). Accounts like these can be found frequently throughout the pedagogical literature (e.g. Abbott, 2002) and are recounted in support of an assumed mnemonic effect of songs. Newham (1995) observed a "close relationship between song and

memory", and Arleo (2000) argued that "the mnemonic value of songs, is, of course, one of the best reasons for using them in the classroom" (p. 10). Empirical research focused on native speakers (L1 speakers) has provided some evidence that song and also poetry can benefit verbal learning. However, findings based on L1-research cannot simply be extended to second language (L2) learners. In the following review I will first recapitulate several pertinent models of human memory relevant for a discussion of a possible mnemonic effect of songs and poems. This is followed by a review of the relevant research on "musical mnemonics" (McElhinney & Annett, 1996) and also poetic mnemonics and a discussion of potential factors rendering songs and poems more memorable than prose texts. Finally, I will discuss in greater detail the small body of existing research on lexical learning through songs in a second or foreign language.

2.3.1 Pertinent models of information processing and storage

A review of models of memory must necessarily be limited in scope and restricted to aspects relevant to the issue at hand, that is, a possible mnemonic effect of songs. First I will outline models that describe memory in terms of a storage system, that is, the Modal Model and the Multicomponent Model. I will then give a brief overview of relevant models that emphasise the processing of information, namely the Levels of Processing Model (LOP), the Transfer-Appropriate Processing Model (TAP) and the Type of Processing-Resource Allocation Model (TOPRA). Finally, I will discuss aspects of long-term memory, particularly the concept of episodic memory.

2.3.1.1 Memory as a storage system: The Modal Model (Atkinson & Shiffrin, 1968) and the Multicomponent Model (Baddeley & Hitch, 1974)

One and possibly the most widely known psychological model of human memory describes memory in terms of storage structures and the information processing that occurs within these structures. Memory is seen as a multi-store system with several separate stores and information flowing back and forth between them. These stores differ regarding how much and for how long they hold information and regarding how and what type of input is processed.

The first and best-known representative of this structure-oriented view is Atkinson and Shiffrin's (1968) Modal Model, which comprises three types of memory stores: a sensory

register, a short-term store or memory and a long-term store or memory. The sensory register is understood as containing "close copies" (Shiffrin, 1999, p. 17) of the information perceived by our senses, for example visual, auditory, haptic, or olfactory information. The short-term memory is a temporary store with a limited capacity of seven plus/minus two items (Miller, 1956) and a limited holding time of roughly 30 seconds (Shiffrin, 1999). However, the retention time can be prolonged by means of holding information in the short-term store, that is, a form of rehearsal or more specifically a control process which is labelled the rehearsal buffer.

The other main memory store is the long-term memory, where information is retained long-term, possibly permanently. Even though the main interest concerning memory is usually the long-term retention of information, the long-term memory is not the main focus of the model. In the Modal Model, the long-term store is a "passive repository" and "an almost unimportant appendage, simply serving as a warehouse for memories" (Shiffrin, 1999, p. 21, 19). Instead, the short-term memory is the crucial active store, both transferring information into and out of long-term memory and also being a workspace or working memory, where all information processing relevant to encoding, storing and retrieval takes place. In Atkinson and Shiffrin's model, the short-term store is thus understood to be the one active system. Critics pointed out that within this model learning is understood to be a function of the time information is held in the short-term memory. However, such a view fails to account for research findings showing that simply holding information in the working memory does not automatically lead to retention. In addition, neuropsychological research (e.g. Shallice & Warrington, 1970) showed that patients with impaired short-term memory are still able to perform cognitive control processes such as reasoning and comprehending and can still form long-term memories.

In response to such challenges to the Modal Model, Baddeley and Hitch (1974) modified and elaborated the concept of short-term memory and proposed the Multicomponent Model, which describes the short-term store as containing several subsystems. In the modified version (Baddeley, 2000), the short-term memory consists of a control system, the so-called central executive, and three dependent so-called slave systems, the visuospatial sketchpad, the phonological loop, and the episodic buffer. These slave

systems perform different functions and use distinct codes to hold information. If one is impaired, the others might still function normally.

The central executive is not a store but a cognitive control system in charge of directing attention and choosing and applying both conscious and automated cognitive strategies, drawing on all three subordinated slave systems. The visuospatial sketchpad temporarily stores and processes visual and spatial input. Of greater importance for verbal learning and for song mnemonics is the phonological loop. The phonological loop temporarily stores and processes verbal input and can be seen as a model of verbal short-term memory (Baddeley, Eysenck, & Anderson, 2009). It comprises a phonological store as well as a phonological rehearsal process, which is "analogous to subvocal speech" (Repovš & Baddeley, 2006, p. 7). The phonological store holds information only temporarily (for a few seconds), and is of limited capacity, that is, it can only hold between five to eight items (cf. Repovš & Baddeley, 2006). Phonological rehearsal can, however, prevent or repair decay. Since rehearsal depends on subvocal articulation, "people can remember as many words as they can say in 2 seconds" (Baddeley, Eysenck, et al., 2009, p. 28). The phonological loop might also be involved in the acquisition of syntax, at least in L1learners (Baddeley, Gathercole, & Papagno, 1998). It has been theorised that syntactic learning is based on the retention of multiword patterns, which then serve as models for speech production and as a basis for the formation of abstract rules. This requires longer multi-word units to be held in the phonological loop.

A fourth component of the current multimodal model is the episodic buffer, a working memory store that can manipulate, integrate and bind information from the other systems in their various distinct codes. The episodic buffer addresses several problems in the previous three-component model (Baddeley, 2000), such as the failure to explain our ability to process long sentences that exceed the phonological loop's capacity. It can retrieve information from the other systems, including the long-term memory, and bind them into larger chunks (cf. Baddeley, 2000; Baddeley, Hitch, & Allen, 2009). Chunking (Miller, 1956) refers to the processing of a string of words (or other items, for example digits) as a larger holistic unit. Both the concepts of subvocal rehearsal and chunking have been associated with a mnemonic effect of songs and will be further discussed below.

2.3.1.2 The processing of information: The Levels of Processing Hypothesis (Craik & Lockhart, 1972)

Craik and Lockhart's (1972) Levels of Processing Hypothesis (LOP) was introduced as an alternative paradigm to the then prevalent focus on the structural side of human memory. Within the LOP Hypothesis, memory is not seen in terms of separate stores, even though the distinction between short-term and long-term memory is accepted as a useful concept (cf. Craik, 2002). Instead, short-term memory is seen as "a temporary activation of parts of [long-term memory]" due to attention being paid to particular "representations that correlate with present experience" (Craik, 2002, p. 307, 308).

The LOP Hypothesis focuses on the processes that occur when incoming information is perceived, processed and ultimately encoded. These processes are understood to be of different quality. On the one hand, there are so-called shallow types of processing that focus on the structural "surface form" (Craik, 2002, p. 308) of words such as the typeface, the number of letters and sound patterns. Such shallow processing is not understood to be conducive to retention. On the other hand there is deeper processing, which refers to semantic and other more demanding cognitive types of analysis, such as visualizing the meaning of a word. These deep processes are understood to result in longer-lasting and much higher rates of retention. Retention is, therefore "a function of depth" (Craik & Lockhart, 1972, p. 676). It is, as Craik and Tulving (1975) put it, "the qualitative nature of the task, the kind of operations carried out on the items that determines retention" (p. 290). However, elaboration, that is, the number and variety of different processes applied, is also crucial to retention. Still, it is claimed that a minimal form of deep processing will result in higher levels of retention than richly elaborated but shallow processing (cf. Craik & Tulving, 1975).

In a series of experiments, Craik and Tulving (1975) found empirical support for the LOP Hypothesis. In most of these experiments, subjects processed words on a shallow, medium or deep level. Shallow processing was operationalised as making a judgment on whether the word was presented in capital or non-capital letters. Medium processing was initiated by asking the subjects to judge whether the test item rhymed with another word. And deep processing was understood to happen when participants were asked to

either judge a word's semantic or syntactic characteristics. In these experiments Craik and Tulving found significantly greater word retention for deep processing.

The Levels of Processing Hypothesis has received some criticism, especially regarding the difficulty to operationalise and measure "depth of knowledge". Nelson (1977), for example, argued that "depth" remains undefined. In his view, the tentative definition provided by Craik and Lockhart (1972) is circular and thus impossible to falsify or verify, as deeper processes are defined by better retention and better retention is seen as a result of deeper processing. Another critical point is the fact that the mode of testing needs to reflect the mode of encoding during the learning event. Subjects perform best when they are tested in the same processing mode as used during learning. Thus, participants engaging in semantic processing of target words will naturally produce higher results on a subsequent test eliciting semantic knowledge than participants who engaged in structural processing. In response to these shortcomings, Morris, Bransford, and Franks (1977) proposed the Model of Transfer-Appropriate Processing (TAP) which is based on the argument that "task meaningfulness must be defined relative to particular learning goals" (p. 519).

2.3.1.3 The Transfer-Appropriate Processing Model (Morris, Bransford, & Franks, 1977)

Morris, Bransford and Franks (1977) conducted a series of experiments which showed that performance on a posttest did not so much depend on the type of processing during learning but rather on the congruence between learning activity and testing task. In other words, when taking a semantic posttest, subjects who focused on the meaning of words during the learning activity outperformed subjects who focused on structural features. However, results were reversed if the test required participants to provide structural rather than semantic information. Accordingly, Morris et al. proposed that the Levels of Processing Hypothesis be modified to incorporate the concept of transfer-appropriate processing (TAP), which understands the quality of a type of information processing to be "relative to the appropriateness of the testing situation" (p. 528).

2.3.1.4 The Type of Processing-Resource Allocation Model (Barcroft, 2002)

In a nutshell, the Model of Transfer-Appropriate Processing (Morris et al., 1977) argues that learners learn what they focus on and should be tested accordingly. Within the field

of second language acquisition research, Barcroft (2002) went one step further: He proposed the TOPRA or Type of Processing-Resource Allocation Model, which describes our processing capacities as limited. When dealing with a demanding task, different processing resources are in competition with each other. As a result, one type of processing will be emphasised at the expense of others. In other words, a learner will focus on meaning at the expense of form and vice versa. As a result, the learning outcome will reflect the dominant processing type. Consequently, a task requiring the learners to process the input with a focus on meaning will show a learning outcome that emphasises meaning at the expense of form. Learners who engage in a task that requires them to process form will do so at the expense of semantic information, and this will be reflected in the information they retain and are able to produce in a subsequent test. In addition, this model implies that semantic processing, which is often understood to be the deeper and therefore more effective form of processing, might not always be appropriate, for example in the early stages of vocabulary learning. When learning new words, Barcroft (2002) found that semantic elaboration can have a detrimental effect on the retention of the formal properties of words. Barcroft (2004), therefore, highlighted "the need to distinguish between different components of L2 knowledge when evaluating how effective a given technique for vocabulary instruction is likely to be" (p. 206).

2.3.1.5 Episodic and semantic memory as part of long-term memory

One important distinction within long-term memory is that of semantic and episodic memory (Tulving, 1972). Semantic memory refers to world knowledge, which includes knowing the meaning of words, knowing concepts, facts, figures, sensory experiences and how to behave in a certain situation. Tulving (1972, p. 386) used the term "mental thesaurus" to illustrate the decontextualised nature of semantic memory. Episodic memory, on the other hand, refers to an autobiographical memory of personally experienced events in the past that includes not only the *what* but also contextual information of the *where* and *when* of the event. While we tend to think of all forms of memory and remembering to be about the past, only episodic memory has, in fact, a direct link to past events (cf. Tulving, 2002). It allows for what Tulving (1985, 2002) has termed mental time travel in subjective time, that is, a re-visiting of a past event.

Neuropsychological research has shown that semantic and episodic memory are indeed two distinct systems. They are, for example, differently affected by attrition. However, research also shows that they are interdependent at the levels of encoding, storage and retrieval. Greenberg, Keane, Ryan, and Verfaellie (2009) found that healthy subjects use episodic memories to produce lists of objects such as kitchen utensils, for example, by imagining their own kitchen and going through all the items they "see". Greenberg and Verfaellie (2010) concluded that, at retrieval, episodic memory often functions as "an organizational strategy or an efficient route to access" (p. 750).

The finding that episodic and semantic memory are separate but interdependent systems led researchers to different conclusions. Tulving (1995) argued for a serial processing of information. In his view, during the encoding phase input is first processed as semantic and then as episodic memory and subsequently stored in both systems in a parallel fashion. Baddeley (1988), on the other hand, drew a different conclusion regarding the interplay of semantic and episodic memory. He saw the two systems as interdependent and argued that semantic memories are formed after a number of (episodic) learning events as an "accumulated residue" (p. 180). In short, semantic memory is decontextualised information abstracted from episodic memory. Tulving (1995) criticised such a view and pointed out that the formation of semantic memories does not require several learning episodes but can be formed during one single event.

2.3.2 Cognitive-psychological research on the potential mnemonic effect of songs and poems

A fair number of studies within the field of cognitive psychology have investigated native speakers' verbatim memory of sung, spoken and rhythmically enhanced spoken texts. The majority of studies have demonstrated a mnemonic benefit of songs compared to spoken texts under certain conditions. However, a small number of studies also produced converse findings. The existing research revealed several factors that can contribute to an increased memorability of songs and also poetic texts.

One of the first studies indicating a facilitative effect of music on verbal learning is Wallace (1994). Wallace conducted four experiments on the retention of folk ballads. She compared different modes of presentation, that is, a sung, spoken, and beat-enhanced

spoken mode and their effect on verbatim text recall and the recall of structural features such as rhyme and the number of syllables. She also explored the effect of the number of verses and the repetition of the melody across verses. In addition, Wallace investigated the effects of one consistent (and hence repeated) melody in comparison to presenting verses with each a different melody. Wallace found that the lyrics of a three-verse ballad were significantly better recalled verbatim if presented in a sung compared to a spoken and beat-enhanced spoken format, on condition that all verses were sung to the same melody. However, when the target text consisted of only one verse, the spoken presentation resulted in higher verbatim recall. Additionally, Wallace observed that the presentation of three verses each sung to a different melody but with the same rhythmic pattern yielded similar results to a presentation in spoken format. Wallace concluded that melody can have a positive effect on recall and reconstruction of text. An acoustically emphasised rhythm alone, however, stressing the same syllables as the melody, seemed to have little facilitative effect. Wallace's study indicates that the melody is a major factor when it comes to the superior retention of song lyrics over spoken text. This finding is supported by a number of other studies.

McElhinney and Annett (1996), for example, found significantly higher free recall of pop song lyrics under a sung learning condition compared to a spoken condition. Chazin and Neuschatz (1990) and Rainey and Larsen (2002) observed increased lexical recall of unconnected lists of words, such as terms for colours and minerals or lists of unknown names of fictitious athletes, following a sung compared to a spoken presentation. And Calvert and Tart (1993) observed superior recall of sung text both in a naturalistic study and in a controlled intervention study. In their naturalistic approach, they investigated whether the frequency of exposure to a sung version of the preamble of the American constitution during childhood in form of a "Schoolhouse Rock" TV vignette affected recall and strategy use during testing. They found that participants who had frequently watched the vignette had a significantly higher recall of the text than participants who had only been infrequently exposed to the broadcast song. In addition, frequent viewers reported a significantly higher use of overt or covert singing as a strategy to remember the lyrics. Calvert and Tart also conducted a laboratory experiment on the recall of the preamble under four different treatment conditions, that is, sung or spoken input

without repetition or presented repeatedly over several weeks. Recall was tested once the treatment was completed and five weeks later. They found that participants of the sung and spoken condition showed similar recall when the text was presented only once. However, in the repeated condition, the song group significantly outperformed the spoken group during immediate posttesting. Consequently, Calvert and Tart emphasised that repetition is essential for melody to serve as a mnemonic aid. This is in line with McElhinney and Annett (1996), who measured free recall after one and several learning trials and found that after only one presentation, test results for the spoken and the sung condition did not differ significantly. While both groups then experienced a significant increase in recall over the next two trials, the song group experienced significantly greater gains.

Wallace (1994) also argued that music constitutes added information that needs to be processed and can therefore be considered an additional memory burden. She argued that repeating a song and its melody a sufficient number of times can transform the melody from a memory burden into a memory aid when it comes to verbatim recall of song lyrics. In a similar vein, Moussard, Bigand, Belleville, and Peretz (2012) found that in patients with Alzheimer's disease "music can interfere with verbal learning in initial stages, and facilitate performance in long-term retention or relearning episodes" (p. 528). Apart from the number of rehearsals, Wallace also found some other factors to moderate the beneficial effect of melody on memory. She qualified that a melody needs to be simple and symmetrical and display a regular rhythm in order to aid retention.

Racette and Peretz (2007), on the other hand, did not find a mnemonic benefit of a sung format, despite repeated line-by-line learning trials and productive retrieval practice. In fact, the authors wondered "why music should facilitate recall, since there is more to learn in a song than in a text" (p. 242). They compared a sung and spoken format not only during learning, but also investigated the effect of singing and speaking during recall. They found that listening to a song and singing did not result in higher verbatim recall of lyrics compared to listening to the spoken text and repeating it in spoken form. They also observed no differences between groups regarding awareness and recall of structural features. Racette and Peretz concluded that "the best strategy for learning lyrics is to ignore the melody" (p. 250). Racette and Peretz also raised the issue of speech rate.

While they controlled for presentation rate, they argued that most studies showing a superior verbatim recall of a sung learning condition did not take into account the slower rate of presentation in songs compared to spoken texts. Kilgour, Jakobson, and Cuddy (2000) looked at the variable of speech rate in even greater detail than Racette and Peretz and found it to be the decisive variable in song mnemonics. Kilgour et al. found a superior memory effect under a sung condition only when the lyrics were presented at a slower rate than in the spoken condition. When presented at a comparable rate, experimental participants engaging with a spoken text produced higher results than the song group by a small but consistent margin.

Another factor contributing to the potential mnemonic effect of songs is the rhythm. While Wallace (1994) found no superior verbatim recall of text presented in a rhythmically enhanced spoken format, Purnell-Webb and Speelman (2008) observed that a spoken learning condition with emphasised rhythm was just as effective in aiding verbatim recall as a sung format, provided that it afforded sufficient rhythmic information. Purnell-Webb and Speelman criticised Wallace's beat-enhanced format for providing only minimal rhythmic information by emphasizing solely the primary accent. In their own study, they employed a more elaborate design which stressed both the primary and secondary accents of words.

A further factor contributing to an increased memorability of songs but also of poems is the poetic language of the lyrics. Tillmann and Dowling (2007) conducted a series of experiments investigating the memory of surface details in poetry and prose. They compared verbatim text retention of poems, which displayed poetic metre and rhyme, and of prose texts. In contrast to the research previously discussed, they did not incorporate a musical condition in their experimental design. However, they understood poetry to be "intermediate between prose and music" (p. 629), sharing semantic content and syntactic structures with prose and resembling music in regard to the temporal organization provided by structural regularities such as rhythm and rhyme. Their testing methods also differed from those employed in other studies. Tillmann and Dowling measured recognition memory rather than recall, using a design of same-different judgments. Participants had to recognise original phrases and identify two types of lures featuring either semantic changes or a changed word order while preserving meaning,

rhythm and rhyme. They found that in the prose condition, correct recognition memory of paraphrased lures decreased over time. However, correct identification rates under the poetry condition even slightly increased over time. Tillmann and Dowling deduced that memory for surface details remained stable when working with poetry, while it seemed to be irrelevant and get lost over time when participants engaged with a prose text. The authors concluded that rhythm and temporal organization, particularly rhyme, improve memory for surface details or at least prevent the loss of such memory over a longer period of time. Dowling, Tillman, and Ayers (2001) showed a similar effect for musical patterns. Tillmann and Dowling speculated that a combination of poetry and music, i.e. song, might prove to be even more effective.

In summary, research with native speakers in the field of cognitive psychology frequently showed superior retention of songs compared to spoken texts. However, a few studies found no difference in verbatim memory between sung and spoken learning conditions. In addition, a small number of studies found that beat-enhanced spoken texts and also poems can result in increased verbatim memory compared to normal spoken texts and prose. Finally, the existing research with L1-speakers implies that among the many elements involved in the potential increased memorability of songs and also poems, the melody, the rhythm and lexical sound patterns are crucial factors contributing to the mnemonic effect of songs. In the following, I will discuss in greater detail how these three factors might affect cognitive processes and increase verbatim retention of lyrics.

2.3.3 The effect of melody, rhythm and lexical structural regularities on memory

The three identified features, melody, rhythm and the structural regularities or sound patterns found in lyrics, can result in cognitive processes benefiting the formation of stable memory traces. Of particular interest is the extent to which text and melody are integrated: Several studies have come to the conclusion that melody and text in songs are closely associated but probably fall short of being stored as a fully integrated holistic unit in long-term memory (see for example Serafine, Crowder, & Repp, 1984; Serafine, Davidson, Crowder, & Repp, 1986; Calvert & Tart, 1993; McElhinney & Annett, 1996; Ginsborg & Sloboda, 2007). However, while many studies propose a beneficial memory effect of melody based on the close association of text and music, Racette and Peretz (2007) spoke of "the cost of singing" (p. 250) and claimed that the melody is, in fact,

detrimental to the recall of lyrics. They argued that the parallel learning of text and melody can lead to a dual-task situation, in which different encoding cues compete for the learner's attention.

Another proposed reason for song's potentially higher memorability is chunking. As explained earlier, chunking is a way of circumventing the limited processing capacity of the phonological loop by binding smaller units of information into larger units, thus effectively increasing the "size" but not the number of units to be processed. As Baddeley (2000) explained, chunking in sentence processing occurs when "additional information, typically from long-term memory, is used to integrate the constituent words into a smaller number of chunks with capacity being set by the number of chunks rather than the number of words" (p. 419). Such additional information is often semantic in nature. However, structural information such as (predictable) patterns of melody, rhythm and rhyme can also encourage chunking beyond sentence level (cf. Wallace, 1994).

Klahr, Chase, and Lovelace (1983) investigated the learning of the alphabet through the well-known "alphabet song". They found that the mnemonic effect of the song is based on its chunking of the letters into larger (rhythmic) units in accordance with the processing capacities of the working memory. In fact, they argued that the song's structure is "a direct result of the properties of the human memory system" (p. 477). In a series of experiments on the retrieval of letters of the English alphabet, Klahr et al. observed that their subjects needed to access the relevant chunk and conduct a serial search within it. In addition, they observed that the formation of the chunks was governed by the learners' expectation of phrases to rhyme. Both Wallace (1994) and McElhinney and Annett (1996) observed indications for increased chunking in experimental participants under the song condition. In contrast, in Purnell-Webb and Speelman's (2008) intervention study neither a melody nor an acoustically enhanced rhythm resulted in significantly different patterns of chunking. Racette and Peretz (2007), on the other hand, did find a certain kind of chunking effect when the recalled text was sung rather than spoken. That is, they found that sung recall was more sequential than spoken recall. However, rather than benefitting text performance, the sequential recall had a detrimental effect, "since recall of a line was more dependent on the recall of the previous line" (p. 248). In addition, Racette and Peretz observed that the forgetting of

one line tended to impede the recall of the following line. This finding is supported by Speelman, Sibma, and MacLachlan (2012), who also found a strong memory for (forward) sequence order in songs.

Another possible effect of presenting language in the format of a song is that melody, rhythm and sound patterns such as rhyme direct the listeners' attention to the form of the lyrics rather than the meaning. When investigating children's recall and comprehension of information presented in songs, Calvert and Billingsley (1998) considered that "the melodic structure could predispose children to listen to how the lyrics sound rather than to what they mean" (p. 106). In what way could such a focus on form aid text retention? This is a relevant question, particularly as it is frequently claimed - for example in the Levels of Processing Hypothesis described above - that semantic elaboration and processing is more likely to lead to transfer of information into long-term memory than so-called shallow processing of surface structures. Wallace (1994) argued that music renders structural aspects of the text more salient to the listener. Wallace and Rubin (1991) elaborated that "rhyme, alliteration, repeating sounds, rhythm, and verse structure help preserve the exact wording" (p. 200) in the transmission of ballads in oral tradition. Rubin (1995) explained that the rhythm can greatly restrict the choice of words when attempting verbatim recall. In addition, structural information can serve as an indication whether all necessary information has been recalled or whether further attempts of retrieval are required (Rubin, Wallace, & Houston, 1993). Purnell-Webb and Speelman (2008) referred to a schematic frame imposed by the rhythm, which serves as a mnemonic device.

Rubin and Wallace (1989) also highlighted that combined cues, in their specific case a combination of semantic cues and rhyme, dramatically limit the choice of possibilities. In addition, they understood such a combination of cues to account for strong differences in research results and general beliefs regarding the effectiveness of rhyme as a mnemonic device:

Rhyme is not an especially good cue for recall in laboratory research (...), yet it is among the most used mnemonic aids outside the laboratory and in oral traditions. (...) The lack of generalization of laboratory results occurs because care

is taken ... to ensure that rhyme and other sound and orthographic properties do not interact with meaning cues. (p. 707)

One could say that research that focuses exclusively on one variable while controlling for other factors as possible interfering variables might obfuscate relevant processes.

Another aspect that needs to be considered is that songs' richness of contextual information encourages the processing and storage of the listening event as an autobiographical or episodic memory, leading to increased recall and recognition compared to spoken texts. Ferreri, Aucouturier, Muthalib, Bigand, and Bugaiska (2013) argued along those lines and proposed that "it is ... possible that the greatest value of music for memory is to provide mnemonic processes with a particularly rich and helpful context during the encoding phase of episodic memory" (p. 2). Research has confirmed the existence of episodic memory for non-verbal music (Palmer, Jungers, & Jusczyk, 2001). And Bartlett and Snelus (1980) argued that participants in a study on very long-term memory for popular songs are likely to have formed episodic memories of the listening events. Songs often evoke a strong emotional response (Krueger, 2011; Krueger, 2014), and emotions can be considered a powerful factor in memory (Sprenger, 1999). Eschrich, Münte, and Altenmüller (2008) found a link between the (positive) valence, or perceived attractiveness, of music and the facilitation of memorization.

Calvert and Tart (1993) proposed involuntary subvocal rehearsal, i.e. the involuntary replaying of a song in one's mind, as a reason for a possible mnemonic effect of melody. Subvocal rehearsal may help retain at least parts of the lyrics in the phonological loop (Baddeley & Hitch, 1974), and memory decay could be prevented or repaired, effectively facilitating the integration of information in the long-term memory. Murphey (1990a, 1990b) referred to this involuntary repetition as the SSIMP, the song-stuck-in-my-head phenomenon. He argued that listening to songs results in a high rate of involuntary subvocal rehearsal of the melody and the lyrics. Apart from a myriad of anecdotal evidence for brainworms (Sacks, 2010), the stuck song syndrome (Levitin, 2006) or involuntary musical imagery (INMI), there is ample empirical support for the existence of auditory imagery, that is the imagining of sound, particularly music, in the majority of people (Halpern & Bartlett, 2011; Williamson et al., 2012; Williamson & Jilka, 2014). Halpern (2001) provided an overview of a number of studies on auditory imagery and

concluded that humans experience both musical and, to a lesser extent, verbal sound imagery in their mind and that such imagined sound is processed at least partially in the same regions of the brain that process actual sound. As Zatorre and Salimpoor (2013) put it: "Think of the first four notes of Beethoven's Fifth Symphony – your cortex is abuzz!" (p. 12).

However, the lyrics in songs have mostly played a subordinate role in research on auditory imagery and have so far often been of interest only as triggers for melody recall in experimental settings. Liikkanen (2012) reported that survey participants experienced earworms both in native and non-native languages, implying that the lyrics were indeed part of the rehearsal experience. On the other hand, in the field of second language acquisition Krashen (1983) proposed the existence of the Din (after Barber, 1980), a purely verbal form of involuntary rehearsal of second language input. He argued that the Din occurs when the language acquisition device (LAD) is being activated as a result of learners receiving comprehensible input.

In summary, research in cognitive psychology has provided evidence that under certain conditions songs and also poems can benefit the retention of verbatim text. The melody, the rhythm and the structural regularities of the poetic text such as rhyme and other sound patterns, have been identified as three main factors rendering songs and also poems more memorable than prose texts. A number of reasons have been proposed as to how these factors affect cognitive processes and consequently benefit the recall and recognition of verbatim text. These reasons include the memorization of text and melody in close association. In addition, it has been argued that the additional structural information provided by the melody, the rhythm and the structural regularities in the language result in the increased processing and encoding of words in larger chunks. Furthermore, it has been argued that melody, rhythm and lexical sound patterns render the linguistic structure of the language more salient and draw the listeners' attention to the linguistic form. Finally, it has been hypothesised that involuntary rehearsal of a song results in increased retention of its lyrics.

The research discussed so far has focused on native speakers. It cannot be assumed that memory processes in second language learners are equivalent to L1-speakers processing

their native language. In addition, studies in cognitive psychology research naturally do not consider the impact of classroom proceedings and teaching techniques. However, in order to measure the effect of songs on lexical learning within the second language classroom, teaching practices need to be taken into account. In the following I will discuss experimental and also classroom-based quasi-experimental studies on lexical learning through songs in second language learning.

2.3.4 Research on a mnemonic effect of songs in L2 studies

Only few empirical studies have investigated the effect of songs on the retention of second language text or vocabulary. Early studies focused exclusively on the use of songs with children. One early empirical study, Hahn (1972), investigated the retention of song lyrics and textbook dialogues by beginning German learners in the seventh grade in the USA. Hahn conducted his quasi-experiment within an educational setting favouring the then popular audio-lingual method, which means that pattern drills, that is, frequent repetition and verbatim memorization of phrases and dialogues, were part of the regular course of instruction. The 38 participants studied a total of four songs and four dialogues over the course of fourteen lessons. Each target text was practiced for 20 minutes during two subsequent lessons. Procedures included a number of activities typical for the audiolingual approach, such as repeated reading of and listening to the target texts, detailed explanations of language and content assisted by visual aids, sing-alongs or read-alongs, the production of written copies of the target texts and retrieval activities. Four posttests measured the retention of previously unknown vocabulary by means of Germancued and an English-cued recognition tests and an uncued and an English-cued recall test. Hahn found that the song condition overall resulted in significantly higher test performance than the spoken-dialogue condition (p < .05). Hahn concluded that songs, while comprising an additional learning burden in the form of the melody, still overall resulted in significantly greater lexical retention. He did not speculate about the reasons for the mnemonic benefit of songs. However, he observed a motivational effect of learning through songs with learners requesting both more time to engage with the songs and copies of the lyrics and the recordings to take home.

Medina (1990) conducted a widely-quoted study that marks the beginning of a greater interest in the empirical study of songs in the language classroom. Medina investigated

incidental vocabulary learning in 48 elementary-school children in the USA. The participants, second-grade students and learners of English as a second language (ESL), engaged with a children's story either in a sung or in a spoken version. In addition, some participants received illustrations of the story as visual aids. The study, thus, included four experimental groups: a spoken/illustrations group, a spoken/no-illustrations group, a music/illustrations group and a music/no-illustrations group. Medina used a multiple choice test with pictorial test items as a pre- and post-test. Since participants only listened to the treatment text, they also provided their answers orally. Medina found no significant differences in vocabulary retention between any of the groups. However, she observed noteworthy patterns in the descriptive data: The highest mean score was obtained by the combined music/illustration group during both immediate and delayed posttesting. And when comparing the two no-illustration conditions, the music group obtained a higher mean score than the no-music group during both immediate and delayed posttesting. Medina concluded that songs produce levels of lexical learning comparable to spoken texts and are therefore an effective means of vocabulary instruction. Given the consistently higher mean scores produced by the music groups when compared to their no-music counterpart, she proposed a possible advantage of songs over spoken texts. Medina speculated that songs in accordance with Krashen's (1989) Input Hypothesis might provide comprehensible input and that their memorability might be enhanced by the music and the rhythm. She also highlighted that the participating children reported preferring the song over the spoken narration.

Hahn (1972) and Medina (1990) both focused on children as participants and revealed equal or superior vocabulary retention following a song learning condition compared to a spoken learning condition. They both used a quasi-experimental approach and introduced additional variables by adding visual aids (Medina) or by measuring the effect of songs in the context of a realistic language lesson involving a variety of activities typical for the then popular teaching method, the audio-lingual method (Hahn). More recent studies have focused on adult second language learners. Smith Salcedo (2002) investigated the effect of a song presentation on cued text recall by 96 university students and beginning-level Spanish learners. The quasi-experimental study involved two experimental groups and a control group. In a counter-balanced design the two

experimental groups read and listened to a total of three popular Spanish songs or to the same texts recorded in the format of a speech. Each song or speech was repeated a fairly high number of six times. During the treatment the learners were assisted in the comprehension of the text. The spoken recordings were produced by native Spanish speakers of the same gender and nationality as the singers. Cued recall was tested immediately and for one song also two weeks later by means of a cloze test. In addition, the possible effect of the melody serving as a retrieval cue was measured: Part of the song group was presented with the melody of the target song while taking the cloze test. Finally, the subjects also responded to a post-treatment questionnaire, reporting on their attitude towards songs in the language classroom, and the occurrence of the Din, that is, the involuntary subvocal rehearsal of treatment text.

The results of the cloze-tests showed a significantly higher immediate text recall in the song group for two of the songs. However, the delayed post-test showed no significant differences between the two learning conditions. Furthermore, listening to the melody while taking the immediate cloze test did not result in superior recall compared to the song group that took the test without listening to a musical cue. In response to the questionnaire, significantly more participants reported having a Din experience under the song condition (67%) compared to the spoken condition (33%), a fact that led Smith Salcedo to conclude that more learning was indeed happening in the song group in comparison to the group listening to the spoken version. The author argued that according to Krashen (1983) the Din is a result of the activation of the language acquisition device (LAD) and thus indicates that acquisition is taking place. However, Smith Salcedo did not address how the lack of a significant difference between learning conditions during delayed posttesting could be explained, in light of the increased Din experience in the song group. Finally, the author also reported that all of the participants responded that songs were a positive addition to the language classroom.

Salas (2006) focused on adolescent learners in a small-scale study that investigated the retention of vocabulary by twelve adolescent ESL learners at two High Schools in New Zealand. In a counter-balanced quasi-experimental intervention study Salas compared vocabulary retention following the presentation of previously unknown target vocabulary in a song or in a poem. In Salas' study three different songs and matching poems were

used, composed specifically for the experiment and the lexical needs of the participants. Apart from the two different learning conditions, Salas introduced a particular teaching technique called audiation technique, an intentional subvocal rehearsal often employed by professional singers to memorise lyrics. Under both learning conditions, participants engaged in this rehearsal practice. The treatment was repeated every day over a period of four days for each song and poem. In addition, participants engaged in other activities and received additional material elaborating on the topic of the particular text. During these activities the target vocabulary was rehearsed. The intervention can thus be described as both intensive and extensive, comprising repeated rehearsal over a fairly long period of time. Salas used a wider range of posttesting than Smith Salcedo (2002), namely active recall, passive recall, active recognition, and passive recognition. The experimental post-tests reveal a significantly higher knowledge gain for the songcondition and Salas tentatively concluded that music can aid verbal retention. However, the study displayed a number of limitations that caused Salas to remark that it should be seen as a pilot study for potential future research. Among these limitations are the reportedly very different teaching styles of the two teachers involved in the study, the limited number of participants, who were also females only, and a great diversity regarding language proficiency among the learners. In addition, some of the additional material and engagement in further practice might have influenced the findings. As Salas pointed out, recognition and recall of target vocabulary relevant for a class excursion to an aquarium was significantly higher for learners participating in the excursion compared to learners who did not join the field trip. It is likely that a greater motivation for learning the target vocabulary and their possible use during the excursion influenced the post-test scores.

The most comprehensive research project to date focused on lexical learning through songs is Ludke (2010). Ludke conducted a series of studies that investigated the effect of songs on verbal learning in a foreign language both in a controlled laboratory setting and in a more natural setting in existing language classes. In the lab experiment Ludke explored the question whether paired-associate phrases in English and in an unknown language (Hungarian) were better remembered when presented in sung format, a spoken format or in a spoken format with emphasised rhythm (using a metronome).

Ludke expected the singing group to outperform the speaking group, particularly on those tests that required oral production of the foreign language, based on Yalch's (1991) findings that differences between learning conditions are predominantly detected when the tasks are particularly challenging. Furthermore, Ludke expected the rhythmic speaking group to fall somewhere between the other two learning conditions, as their treatment comprised only one of the musical elements of the singing condition, that is, the rhythm.

Sixty university students listened to 20 expressions in English and Hungarian repeatedly for 15 minutes. In addition, during the first 5-minute cycle of phrase presentation, the expressions were also visually displayed on a computer screen. Furthermore, participants were encouraged and were given ample time to orally or subvocally rehearse the phrases. Ludke tested various aspects of productive and receptive knowledge of the paired English and Hungarian target phrases in a series of four tests immediately following the treatment. In addition, she used a delayed productive Hungarian conversation test 20 minutes after the end of the treatment, during which participants were prompted with a Hungarian phrase and were asked to produce an appropriate response in Hungarian.

Ludke found that the song condition resulted in the highest mean scores on four of the five tests. Only on a recognition test did the speaking group reach a similar mean score as the singing group. A series of analyses of variance for the five tests revealed a significant main effect for learning condition only in the two more demanding tests of immediate Hungarian production (cued by the English associate) and the delayed conversation test, confirming Ludke's previously stated predictions. A pairwise post hoc analysis revealed that on the Hungarian production test the singing and speaking condition differed significantly with the singing group clearly outperforming the speaking group and the rhythmic-speaking group falling between the two. On the conversation test the singing group reached a significantly higher mean score than the rhythmic group.

Interestingly, Ludke also measured the effect of a number of individual learner differences on verbal learning and found that the extent of previous language learning

experience and a positive mood at the beginning of the overall experimental session had an even greater influence on test performance than the learning condition.

Ludke conducted a second experiment, which compared a singing and a speaking condition only and left out the rhythmic-speaking condition. In addition, she added the variable of active oral rehearsal (listen-and-repeat) versus a passive receptive learning condition (listen-only). Surprisingly, in this experiment the speaking group consistently scored higher than the singing group and, in fact, reached a significantly higher mean score on the previously innocuous recognition test. Also contrary to expectations, the passive listen-only group produced significantly higher results overall and specifically on the English recall test. Ludke argued that this surprising contradiction could at least partially be explained by pre-existing individual learner differences, as the participants in the two experimental groups were not well-matched regarding particular relevant characteristics. The singing group, in particular, had significantly lower musical production abilities, which is unfortunate given the fact that participants in the listenand-repeat condition were required to sing along with the presented target phrases. In addition, they displayed significantly lower motivation scores, a factor that had previously proven to be a relevant variable affecting test performance. Ludke concluded that singing can lead to improved learning of phrases if participants are well matched for individual differences.

In addition to the highly controlled laboratory experiments, Ludke (2010) also conducted a classroom-based quasi-experimental intervention study with two existing classes of early-teenage beginning learners of French in Scotland. Over the course of four weeks, each class received six 15-minute lessons of a music intervention involving a song and six lessons of an arts intervention involving a dramatic dialogue. Both groups participated in both interventions, however in reversed order and with different songs and dialogues. Ludke described the intervention as "listen-and-repeat learning" incorporating elements of the Contemporary Music Approach (Anton, 1990) and recommendations made by Welch (2007) regarding the teaching of choral music in a foreign language. The lessons involved a wide range of activities such as listening, drawing, a cloze exercise, comprehension questions in French, the learners' production of their own comprehension questions, practicing a choral performance of the song or dialogue,

producing an audio recording of the performance and repeated listening to the recorded performance. Ludke conducted a series of pre-, mid-point-, and post-tests during the treatment including a translation test of phrases taken from the target songs and dialogues, a cloze test based on the original texts, and a student questionnaire.

Overall, Ludke found no superior test performance following the music intervention. One group, in fact, experienced a decrease in test performance following the music intervention. Ludke speculated on the reasons for this decrease which occurred during the second half of the intervention study. She hypothesised that the learners might have displayed a drop in motivation regarding participation in the study as well as an increase in stress due to a French exam taking place at the time. In addition, the researcher pointed out that the song might have been markedly more difficult than the dramatic dialogue. Ludke concluded that both music and artistic arts interventions can be considered beneficial activities for beginning French learners. Finally, the student questionnaire revealed that learners perceived an increase in self-confidence in speaking the target language. In addition, the survey showed that a majority of learners in both classes (62.5%) preferred the song to the dialogue. A majority also reported experiencing the Din after listening to the song (52%), while clearly fewer students experienced this phenomenon after engaging with the dialogue (9.4%). Ludke concluded that given the high preference of learners for songs and the overall positive learning outcome, songs should be incorporated in modern language curricula.

As a review of the existing L2-research reveals, no study has taken into account structural regularities, that is, sound patterns, as a factor rendering songs mnemonic. While Ludke incorporated a beat-enhanced and therefore rhythmic learning condition, no study investigated the effect of a poetic text, that is, a text comprising poetic features such as alliteration, assonance and end-rhyme, on the retention of words, phrases or longer connected text in comparison to a prose text. There is however some evidence that L2-learners can benefit from sound patterns when learning word strings, particularly when these patterns are explicitly brought to their attention. For example, research has shown that EFL learners show greater recall of alliterative words strings than non-alliterative word strings, both when studied as decontextualised items (Boers & Lindstromberg, 2005; Lindstromberg, Boers, & Eyckmans, 2014) or under incidental learning conditions

when encountered in longer texts while primarily focusing on content (Boers, Lindstromberg, & Webb, in press). Assonance has also been found to benefit the memorization of multi-word units when encountered in a context that promoted the noticing of form, albeit to a lesser degree than the seemingly more salient phonemic pattern of alliteration (Lindstromberg & Boers, 2008).

In summary, only few empirical studies have investigated second language learners' and particularly adult learners' lexical learning through songs. While several studies revealed a positive effect of songs on text and phrase retention (Hahn, 1972; Smith Salcedo, 2002; Salas, 2006; Ludke, 2010), others found no difference between a sung and a spoken learning condition (Medina, 1990; Ludke, 2010 in her classroom-based study). With the exception of Ludke's (2010) laboratory experiments, all studies have taken a classroombased, quasi-experimental approach. As a result, they incorporated a number of additional variables into the design of the intervention. These additional factors include the number of repetitions of the target texts, massed and distributed rehearsal, the addition of visual aids and of a whole range of complementary classroom activities and auxiliary teaching techniques.

While it is appropriate to measure the success of a musical teaching approach in a realistic classroom setting, none of the existing studies have assessed the extent to which the design of their treatment overall reflects actual teaching practices. In some instances, it can be assumed that the intervention procedures differ from the current teaching approach involving songs. In her classroom-based study Ludke (2010), for example, incorporated four-weeks-long arts and music interventions with an array of fairly elaborate activities. Given the high investment of time and effort, this could be considered an uncommon approach. Salas (2006) integrated the *audiation technique* into her design, a rehearsal method common among professional singers but not among language teachers. In addition, Salas used songs specifically written and composed for the study, and it can be speculated that few teachers have the musical training and the time to do so. Medina (1990) added visual aids as a variable, and Hahn (1972) applied an elaborate treatment involving a number of activities. While his intervention seems to have reflected procedures typical for the audio-lingual method, they might not be equally reflective of the currently popular approach of communicative language teaching.

Finally, L1-research in cognitive psychology has provided evidence that at least three factors, the melody, the rhythm and sound patterns found in the poetic language, are crucial to the mnemonic potential of songs. Only Ludke (2010) in her laboratory study incorporated a rhythmically spoken learning condition in her experimental design, thus taking into account that variables other than the melody might affect verbatim text retention.

The present quasi-experimental study was designed to add a further angle to the small body of research on lexical learning through songs in adult second language learners. It prioritised ecological validity; that is, its intervention was designed to reflect current teaching practices as much as possible. This was achieved by integrating information about typical teaching practices, including the choice of material, into the design of its intervention. As previously discussed, this information was elicited by means of a teacher questionnaire and a song corpus study. Furthermore, it compared three rather than two experimental groups by including a poem condition in addition to a song condition and a spoken prose learning condition.

2.4 A summary of the literature review

This chapter reviewed literature relevant to three different aspects of song-use for L2teaching. Firstly, it looked at existing research into actual teaching practices and teacher cognitions. Secondly, it reviewed literature relevant to a discussion of the lexical profile of songs used in L2-teaching. This included a discussion of a frequent methodology to assess vocabulary load and the research on the lexical profile of other authentic English text genres used in language teaching. Finally, the review discussed the potential mnemonic effect of songs and reviewed research on word and text retention in second language learning. The first section of the review showed that few studies have investigated how songs are actually used by teachers inside the language classroom. These studies all looked at very small groups of very specific informants (for example in terms of nationality and institution). In addition, these studies were mostly concerned with beliefs and attitudes and did not elicit much detailed information about how songs are actually used. In response to these limitations, Chapter 3 presents a survey study based on an international online teacher questionnaire involving over 500 respondents.

This questionnaire elicited a great many details on the actual use of songs in language classrooms around the globe.

The second part of the literature review revealed that we know much about the vocabulary load of a variety of text genres but that we know little about the lexical profile of songs used as teaching material. Chapter 4 presents the compilation and analysis of a song corpus focused on the lexical demand and vocabulary learning potential of songs used as material. In addition, the corpus study compared the lexical load of songs to that of other genres.

The third part of the literature review revealed that research in the field of cognitive psychology repeatedly observed superior verbatim retention of song lyrics compared to spoken texts by native speakers. The melody, the rhythm and structural regularities of the lyrics are frequently pointed out as factors contributing to a mnemonic effect of songs. However, research focused on L2-lexical learning through songs in general and on second language learners' verbatim memory of sung and spoken words, phrases and texts in particular are scarce. The review further highlighted that no previous study has investigated whether songs as teachers actually tend to use them benefit lexical learning. Chapter 5, consequently, presents a study on the mnemonic effect of songs compared to poems and prose texts. The study emphasised ecological validity, that is, the quasi-experimental design was informed by actual teaching practices and materials (as described by the teacher questionnaire and the corpus study).
Chapter 3 The teacher survey

3.0 Introduction

The survey was designed as an online questionnaire asking second and foreign language teachers if, how and why (not) they used songs in the classroom, with a particular focus on vocabulary teaching. Figure 3.1 provides an overview of the main characteristics and purposes of the survey. The complete questionnaire can be found in Appendix A.

Goal	Exploratory investigation of teachers' use of songs in the classroom, with a particular focus on vocabulary teaching;
Format	International online questionnaire (Qualtrics®);
Informants	568 second/foreign language instructors, particularly ESL/EFL teachers;
Distribution	Snowball distribution; advertising in pedagogical online newsletter in NZ, requests through mailing lists and discussion boards of several teacher associations; personal and professional contacts;
Number of questions	Total: 43; minimum: 26;
Question types	Questions eliciting demographic, attitudinal and behavioural data; 33 closed, 8 semi-open ("other option") and 2 open questions;
Link to corpus	The survey provided EFL/ESL songs for the compilation of a song corpus;
Link to experiment	The survey provided information about the typical use of songs in the L2-classroom (ecological validity); it also provided an understanding of the reasoning behind teachers' song use, which influenced methodological decisions;

Figure 3.1: An overview of the teacher survey

In the following, I state the eight fairly broad research questions of this exploratory study. In the subsequent methodology section, I first detail the design, piloting and distribution of the questionnaire. I then describe in detail the informants who participated in this survey study. This description includes data about the respondents' gender, work environment and cultural and educational background. The informants' personal information is also later analysed regarding a potential correlation with teaching practices involving songs. The methodology section is followed by a description of the data analysis, by a presentation of the results and a discussing of the main findings of the study contributing to the overall research question of this thesis. Following the discussion, I point out the particular limitations of the present survey study.

3.1 Research questions

The questionnaire was designed to answer the following questions:

- Do language teachers consider songs to be a useful teaching tool and do they use songs in the classroom?
- 2. For what purposes, if any, do teachers use songs in the classroom?
- 3. If language teachers do not use songs in the classroom: why not?
- 4. What personal, institutional and societal factors influence teachers' use of songs?
- 5. Do teachers use songs to introduce or practise vocabulary?
- 6. According to teachers, what characteristics, if any, make songs suitable for vocabulary acquisition?
- 7. How do language teachers use songs in the classroom?
- 8. What specific songs do EFL/ESL teachers use for what reasons, to what ends and for what levels?

3.2 Methodology

The questionnaire was designed using the survey software Qualtrics[®] (2011) and was informed by scholarship in the area of survey design, particularly by Bradburn, Wansink and Sudman (2004), Dörnyei (2007) and Dörnyei and Taguchi (2010). The electronic survey tool allowed for the use of so-called skip logic. That is, if respondents selected a particular response to certain items they consequently skipped some of the following items. For example, informants stating that they did not use songs in the classroom skipped all items requesting information about the details of their song-use while still being presented with other relevant questions. The questionnaire was divided into three main parts, which elicited information about (I) the informants' opinions about songs as teaching tools and their teaching practice involving songs, (II) the informants' demographic data and (III) the informants' general interest and expertise in music. Part I also contained a sub-section with a particular focus on vocabulary teaching. In addition, it contained one item requesting song recommendations (item 9). English songs provided by EFL/ESL teachers were later used to compile the pedagogical song corpus discussed in Chapter 4. Figure 3.2 provides an overview of the questionnaire and its different sections.

SECTION	SUB-SECTION QUESTION		NO. OF QUESTIONS
 Information and conser A quick overview of the Some technical informa Currently teaching? 	nt questionnaire tion		
Preliminary question	What language do you teach?	Q01	01
I. Your opinion and teaching practice	Ia. Songs in the classroom	Q02-Q19	19*
	Ib. Songs and vocabulary	Q20-Q22	05*
II. A bit about you		Q23-Q31	11
III. Your musical interests and training	IIIa. General musical interest and training	Q32-Q35	04
	IIIb. Own experience as language learner	Q36-Q37	03*
**	Total/maximur	n/minimum:	43/42/26

Figure 3.2: A breakdown of the questionnaire

*Depending on the response to particular items, not all questions were displayed.

Answer options came in various formats such as semantic scales, multiple-choice lists, dichotomous yes/no-answers and open-answer formats.

3.2.1 Survey design and piloting

Initially, 46 survey items were drafted. Their development was informed by a literature review (see Chapter 2), by informal discussions with language teachers, and by an informal review of online teacher discussion boards as well as of lesson plans offered on pedagogical websites and in textbooks. In particular, this broad assessment of teacher opinions and teaching materials informed the creation of prefabricated answer options, that is, answers that were predetermined during the questionnaire design and required

informants to select those responses that they agreed with. While such pre-fabricated answers will necessarily influence the informants' response pattern, they were provided to ease the burden on the informants and to avoid an even higher drop-out rate than could be expected considering the length of the questionnaire. An "other-option" was included for informants to provide additional, unanticipated information. Open-ended answers were analysed and either incorporated into existing categories or used to establish new categories. The initial set of items underwent a critical review by my supervisors and a statistician. A revised set was then given to 21 language teachers and/or applied-linguistics researchers for critical review. This stage included a guided critical discussion of the survey with 15 of the reviewers. As a result, several changes were made before a pilot survey was launched. These changes affected the order of sections and the location of item 9 (song recommendations). Also, the review process led to the addition of a preliminary question asking whether informants were "currently teaching" or not (see below). The responses of ten pilot informants were analysed, and further changes were made. The survey then received a final layout make-over in order to make it as user-friendly and accessible as possible. The final version was launched on March 16, 2011 and was accessible to informants for four months.

3.2.2 Ethical considerations

The survey was approved by the Human Ethics Committee at Victoria University of Wellington on January 21, 2011. All documentation can be found in Appendix A.

3.2.3 Distribution

The invitation to participate in the survey was undertaken in a non-randomised manner through networking based on professional and personal contacts within the field of language teaching and research, through online discussion boards and mailing lists of several national and international language teacher associations and through an advertisement in the online newsletter of the New Zealand national association of teachers and tutors of ESOL (TESOLANZ). Distribution relied on snow-balling (Dörnyei & Taguchi, 2010). In other words, informants were asked to distribute the questionnaire further among colleagues and friends. It has to be acknowledged that this nonrandomised selection process is likely to have impacted on the data elicited. In particular,

despite the repeated declaration in the title and the survey description that the questionnaire seeks insights from both "users" and "non-users" of songs, the mode of selection may have attracted a majority of teachers who take a positive stance on the use of songs for pedagogical purposes. That is, song-users might have been more inclined to participate in research on a subject that they take a decided interest in. They might have understood the survey to back songs as teaching tools and to support the method's propagation in the teaching community. As a result, the ratio of users and non-users in the survey data may not provide an accurate picture of the actual ratio among language teachers in general.

3.2.4 Informants

The international survey was targeted at male and female second and foreign language (L2) instructors of all languages with diverse ethnic, cultural and educational backgrounds in a range of countries working at both private and public institutions and as selfemployed private teachers of adults and adolescents. A total of 568 informants gave initial consent for their response data to be used for research purposes. However, only 398 or 70.1% completed the questionnaire. That equals a total drop-out rate of 29.9% over the course of the questionnaire. Informants dropped out at all points. Each item, thus, was answered by varying and gradually diminishing total numbers of informants. The number of informants responding to different items also varied due to the fact that not all items were presented to all respondents. Questions regarding the use of songs, for example, were not displayed to respondents who initially stated that they did not use songs. Inferential statistics could, therefore, only be undertaken for those informants responding to the correlation.

3.2.4.1 The definition of "language teacher"

One issue that was raised during the pilot phase was the definition of who, in fact, could be considered a language teacher. Many language instructors, particularly self-employed teachers, often find themselves between jobs. In addition, many teachers interrupt their teaching career, for example, to get an advanced university degree, before resuming their work as language instructors. It was, therefore, necessary to point out to (potential) respondents that both currently active teachers as well as people who consider

themselves teachers but are not currently engaged in language teaching were considered viable informants. A preliminary item in the questionnaire was created to clarify this issue: "Please indicate below whether you are currently teaching or not. If you consider yourself to be a language teacher but you are not currently teaching, please think of the last time you were actively involved in language instruction and answer the questions accordingly." Of 523 informants, 81.6%, that is 427, were "currently teaching", while 18.4% (96) were "not currently teaching".

3.2.4.2 The informants' target languages

514 informants taught overall 20 different languages. However, the vast majority (390 or 76%) taught English. This result is due to the distribution of the questionnaire primarily to English teachers through EFL/ESL mailing lists and through professional contacts in the field of teaching English to speakers of other languages (TESOL). Other comparatively frequent target languages were Japanese, German, French and Spanish. Table 3.1 shows the exact numbers and percentages of the reported target languages. In the survey design, I anticipated a potential problem with instructors teaching more than one language, as teaching practice might differ for different languages. For example, an informant teaching both English and Latin – not an unusual combination for German High School teachers – might use songs for the former but not for the latter. I therefore clarified: "If you teach more than one language, please choose one and answer all questions thinking of that particular language."

Target language	Ν	%
English	390	75.9
Japanese	32	6.2
German	28	5.4
French	23	4.5
Spanish	15	2.9
Arabic	6	1.2
Maori	4	0.8
Italian	3	0.6
Chinese	2	0.4
Farsi, Greek, Indonesian, Latin,		
Latvian, Portuguese, Russian, Samoan,	1	0.2
Thai, Ukrainian, Urdu <i>(each)</i>		

Table 3.1:	Languages	taught by	the	informants
------------	-----------	-----------	-----	------------

3.2.4.3 The informants' gender and age

Of 405 informants, 31% (126) were male, 69% (279) were female. The average age of 402 participants was 41.0 years (with a standard deviation of 11.9 years). The youngest informant was 22, the oldest was 87.

3.2.4.4 Current institution

Respondents reported teaching at a variety of different types of institutions. Since many teachers often work in several employment situations simultaneously, this survey item allowed for multiple selections. Table 3.2 shows that the majority of 402 informants worked at a tertiary institution (university, tertiary college, polytechnic), followed by employment at a secondary school (which includes secondary vocational schools) and a public or private language school (which includes heritage language schools). Between ten and twenty per cent of informants were employed in continuing or adult education (including language and life skill courses for immigrants) or as self-employed private teachers (which includes online lessons and courses offered for businesses and companies). Only twelve respondents reported teaching at post-secondary vocational schools. Eleven informants were teaching at primary schools. The fact that 38 informants worked with children rather than adolescents or adults is a limitation of this study as my original intention was to focus on the (young) adult classroom. This limitation is further discussed below.

Type of institution	Ν	%
Tertiary education	189	47.0
Secondary school	98	24.4
Language school	83	20.6
Continuing education	59	14.7
Self-employed	49	12.2
Primary school	38	9.5
Vocational training	12	3.0
Currently not teaching	11	2.7

Table 3.2: The ir	formants'	institution	(s)
-------------------	-----------	-------------	----	---

3.2.4.5 Nationality and current country

402 informants claimed a total of 53 different nationalities. Due to the manner of distribution of the questionnaire and the prevalence of ESL/EFL teachers, most participants were US-American, New Zealanders, Canadian, British, Malaysian, Japanese, German or Australian. Table 3.3 provides exact numbers for the informants' reported nationalities.

Nationality	Ν	%
American	66	16.42
New Zealander	61	15.17
Canadian	50	12.44
Malaysian	31	7.71
British	31	7.71
Japanese	27	6.72
German	21	5.22
Australian	13	3.23
Spanish	9	2.24
Indian	8	1.99
Chinese	7	1.74
Brazilian; Iranian; Uruguayan <i>(each)</i>	5	1.24
Venezuelan	4	1.00
Belgian; Egyptian; French; Greek; Italian; Saudi; <i>(each)</i>	3	0.75
Algerian; Colombian; Indonesian; Irish; Pakistani;	р	0.50
Sudanese; Taiwanese; Ukrainian; Vietnamese (each)	2	0.30
Austrian; Bulgarian; Chilean; Comoran; Dutch; Filipino;		
Iraqi; Jordanian; Kazakhstani; Latvian; Lebanese;	1	0.25
Mexican; Moroccan; Omani; Polish; Romanian; Russian;		
Somali; Swiss; Syrian; Thai; Tunisian; Turkish (each)		

Table 2 2.	Tho ir	formants'	nationalities
TUDIE 3.3.	пе п	ijormants	nationalities

The informants were located in 41 different countries, with the largest group of informants teaching in New Zealand. Comparatively large numbers of informants were also located in Japan, Canada, the United Arabic Emirates (UAE), Malaysia and the USA. Table 3.4 gives an overview of how many respondents taught in the reported countries.

Country	Ν	%
New Zealand	81	20.15
Japan	51	12.69
Canada	49	12.20
United Arabic Emirates	29	7.21
Malaysia	28	6.97
USA	27	6.72
Germany	19	4.73
Australia	16	3.98
Spain	10	2.49
Saudi Arabia	10	2.49
South Korea	9	2.24
China	8	1.99
Italy	7	1.74
Uruguay	6	1.49
Oman; Taiwan; United Kingdom; Venezuela (each)	4	1.00
Belgium; Indonesia; Iran; Vietnam (each)	3	0.75
Algeria; Egypt; Greece; Jordan; Qatar (each)	2	0.50
Colombia; Comoros; Czech Republic; India;		
Kazakhstan; Kenya; Lebanon; Macedonia; Mexico;	1	0.25
Poland; Sudan; Turkey; Ukraine; Yemen (each)		

Table 3.4: The informants' geographical location

3.2.4.6 Years of teaching experience

The majority of 401 informants were highly experienced in language teaching: 50% or 200 reported having 11 or more years of teaching experience. 21.4% (86) reported 6 to 10 years of teaching experience. 17% (68) had 3 to 5 years, while only 7.7% (31) had just 1 to 2 years and only 4% (16) had less than one year of experience.

3.2.4.7 The informants' teaching experience with particular proficiency levels

When it came to teaching different proficiency levels, 400 informants could make multiple selections, since teachers are likely to teach at more than one level. Table 3.5 provides an overview of how many teachers reported having experience teaching particular proficiency levels. Most informants claimed experience with low- and highintermediate learners and continuing beginners. Fewer respondents reported having worked with low-advanced learners and absolute beginners. A considerably smaller number claimed experience teaching high-advanced learners. This is not surprising, as it appears that there is often much less demand in language teaching at highly advanced levels.

Proficiency level	Ν	%
Absolute beginners	263	65.8
Continuing beginners	321	80.3
Low-intermediate	246	06 5
learners	346	86.5
High-intermediate	222	00.0
learners	323	80.8
Low-advanced learners	265	66.3
High-advanced learners	168	42

Table 3.5: The informants' teaching experience with particular proficiency levels

3.2.4.8 The informants' teaching experience with particular class types

400 informants had experience teaching a wide range of class types. Again, multiple answer selections were possible. Table 3.6 gives an overview of the informants' reported experience with various class-types. Over or close to 60% claimed experience with foreign and second language courses or courses focusing on conversation, culture or literature. Slightly fewer informants reported having experience teaching courses for academic purposes, whereas only 38.5% reported experience teaching courses for specific purposes, that is, specific professional or academic fields such as business, tourism or law, for example English for specific purposes (ESP). A small number of informants made use of the "other option" and reported additional experience with teaching specialised courses such as exam preparation (11 informants), teacher education and development (7), literacy and life skills for immigrants (4) and translation and interpretation (2).

Class type	N	%
Foreign language course	257	64.3
Second language courses	242	60.5
Conversation, culture & literature	234	58.5
Academic purposes	214	53.5
Specific purposes	154	38.5
Exam preparation	11	2.8
Teacher education	7	1.8
Literacy & Life	4	1
Translation	2	0.5

Table 3.6: The informants' teaching experience with particular class types

3.2.4.9 The informants' interest and expertise in music

The personal interest as well as expertise in music was also decided to be important for an adequate description of informants. A vast majority of 399 respondents, that is 85.7%, reported that music was very important or important in their personal life.

Response	N	%
Important/very important	342	85.7
Neither, nor	44	11.0
Not very/not at all important	13	3.3

Table 3.7: Importance of music in private life

A majority (81%) of the 399 informants listened to vocal music very often or often in their private life. Non-vocal, that is, instrumental, music was not as popular. Still, around 50% of informants listened to it often or very often.

	Vocal music		Non-voc	al music
Frequency	N	%	N	%
Very often	189	47.4	97	24.3
Often	135	33.8	107	26.8
Sometimes	69	17.3	163	40.9
Never	6	1.5	32	8.0

Table 3.8: Informants' listening habits (vocal and non-vocal music)

Many informants reported at least some expertise in making music, particularly instrumental music. Of 399 informants, 70% or 281 played or had been playing an

instrument from anywhere between less than one to more than eleven years. 61% or 244 had anywhere between less than one to more than eleven years of practice singing in a vocal ensemble.



Table 3.9: Informants (in percent) singing in a vocal ensemble (black)or playing an instrument (grey)

Another way of assessing musical expertise was to ask whether informants were able to read sheet music. Roughly equal numbers of 399 informants were able (48.4%) or unable (51.6%) to read music.

3.3 Data analysis

Response data for each survey item was collected within the survey software Qualtrics[®] and exported to IBM SPSS Statistics, Version 19 (2010) as well as Microsoft Excel. SPSS was used to conduct a number of statistical analyses. In order to test for possible associations between particular items, for example between the informants' gender and their use of songs, a Pearson chi-square test of independence was conducted. However, in several cases the data was clearly skewed and a linear by linear association chi-square test or a simple comparison of descriptive statistics was more appropriate.

3.4 Results

The presentation of the questionnaire results is organised in a way that follows the order of research questions, answering each question in turn. Firstly, I present the informants' views regarding the usefulness of songs for language teaching and their actual use of songs. Secondly, I report on the respondents' purposes for utilizing songs in class. Thirdly, I investigate why some informants did not use songs in the classroom. I then present various personal, demographic, institutional and societal factors that might have

influenced the informants' teaching practice. The subsequent section focuses on teaching vocabulary. Furthermore, I report in detail how exactly informants reported using songs in the classroom, including the number of repetitions, the time spent on a song-based teaching unit and the use of complimentary activities. Finally, I take a brief look at the song recommendations made by the survey informants. As the recommended songs formed part of the pedagogical song corpus of Study 2, they are discussed in greater detail in Chapter 4. In reporting these study results I sought to give informants a voice beyond the mere response statistics. To do so, I have incorporated a number of quotations, which highlight a particular point or provide additional detail. Such quotations can be found throughout the section and are signposted by quotation marks.

3.4.1 The informants' views on songs' usefulness and their actual use of songs in class

The first question, the survey intended to answer was: Do language teachers consider songs to be a useful teaching tool and do they actually use songs in the classroom? Initial questionnaire items assessed the informants' opinions regarding the usefulness of songs in general and specifically for different proficiency levels and class types.

3.4.1.1 Are songs a useful tool in the language classroom?

A large majority, that is, 87.5%, of informants agreed or strongly agreed with the statement that "songs are a useful tool in the language classroom to foster language acquisition". As Table 3.10 shows, only a very small number of informants actually (strongly) disagreed with this statement.

Response	Ν	%
Strongly agree	204	39.9
Agree	243	47.6
Neither, nor	52	10.2
Disagree	9	1.8
Strongly disagree	3	0.6

Table 3.10: Usefulness of songs to foster language acquisition

Usefulness for different proficiency levels

495 informants provided their views on the usefulness of songs for learners of different proficiency levels. Table 3.11 gives an overview of the informants' responses.

Response % (N)	Absolute	Continuing	Low-	High-	Low-	High-
	Beginners	Beginners	intermed.	intermed.	advanced	advanced
Agree/strongly	81.7	87.6	87.3	79.2	73.3	61.6
agree	(404)	(434)	(432)	(392)	(363)	(305)
Neither, nor	9.7	9.1	10.3	16.2	18	24.2
	(48)	(45)	(51)	(80)	(89)	(120)
Disagree/strongly	8.7	4	2.4	4.6	8.7	14.1
disagree	(43)	(16)	(12)	(23)	(43)	(70)

Table 3.11: Usefulness of songs when teaching different proficiency levels

Overall, informants considered songs to be useful for all levels. The strongest support was voiced for using songs with continuing beginners and low-intermediate beginners with almost 90% agreement. Around 80% of informants found songs useful when working with absolute beginners and high-intermediate learners. Agreement was lowest for advanced learners, particularly for high-advanced learners.

Usefulness for different class types

475 informants rated the usefulness of songs when teaching the following class types: second language courses for general, everyday purposes (for example ESL), foreign language courses for general, everyday purposes (for example EFL), language courses for academic purposes, i.e. for study purposes in higher education (for example EAP), language courses for specific purposes, i.e. for specific professional or academic fields such as business, tourism or law (for example ESP), and conversation classes. Table 3.12 shows the agreement of informants regarding the usefulness of songs when teaching different classtypes in percent and in total numbers.

Response % (N)	General SL	General FL	Conversation	Academic	Specific
Agree/strongly	88.9	87.8	84.2	39.9	34.9
agree	(423)	(418)	(401)	(190)	(166)
Neither, nor	7.8	9.2	9.9	34.2	34.5
	(37)	(44)	(47)	(163)	(164)
Disagree/strongly	3.4	2.9	5.9	25.8	30.7
disagree	(16)	(14)	(28)	(123)	(146)

Table 3.12: Usefulness of songs when teaching different class types

Songs were widely seen as useful tools for more general language courses, that is, for second and foreign language courses and for conversation classes. Using songs in language courses for academic and specific academic or professional purposes, however, found clearly less support among informants.

3.4.1.2 The informants' use of songs

The previous items established that a large majority of informants considered songs to be useful when teaching. The next question inquired whether they actually used songs in the language classroom. Results show that a clear majority of informants did use songs in the classroom. Of 476 informants, 391 or 82.1% used songs, while 85 or 17.9% did not. Interestingly, these numbers are not in accord with the ratios previously displayed when informants provided their views on the general usefulness of songs to foster language acquisition. Neither the number of song-users nor the number of non-users reflects the earlier ratio of approval and disapproval of song use. It is unlikely that this mismatch is simply due to varying total numbers of informants, which becomes particularly apparent when comparing the number of non-users and previous disapproval rates: 85 informants reported not using songs, but only 12 informants previously reported that they considered songs to be inappropriate for language teaching. Apparently, factors other than a negative opinion must be influencing the non-users' teaching practices. Section 3.4.3 below examines why some informants do not use songs, and section 3.4.4 explores various additional factors that might influence the respondents' choices.

The informants' use of songs with learners of different proficiency levels

As we saw earlier, informants found songs to be useful for learners of all proficiency levels, with larger numbers of respondents agreeing when teaching continuing beginners and low-intermediate learners and a smaller – but still considerable – number of respondents agreeing when teaching high-advanced learners. However, as the previous item showed, a positive opinion did not necessarily translate into an actual use of songs in class. An additional questionnaire item, therefore, explored the use of songs with various proficiency levels. As teachers typically do not teach all levels, they had the option to choose the answer "not applicable". Informants who selected this option were not included in the calculations. Table 3.13 shows the respondents' use of songs with

different proficiency levels, arranged from most to least frequent. Results show that the use of songs with learners of different proficiency levels generally reflects the informants' opinions expressed before: Songs are used at all levels, but most commonly when teaching low-intermediate learners and continuing beginners. Fewer but still a significant number of informants use songs with advanced learners. Differences in the total percentages compared to previous items can be explained by the fact that only informants who actually taught these levels were considered.

PROFICIENCY	Yes (N)	Yes (%)	No (N)	No (%)	Total applicable
Low-intermediate	335	94.6	19	5.4	354
Continuing beginners	311	93.7	21	6.3	332
Absolute beginners	245	85.7	41	14.3	286
High-intermediate	274	82.8	57	17.2	331
Low-advanced	215	76.5	66	23.5	281
High-advanced	148	64.6	81	35.4	229

Table 3.13: Use of songs with learners of different proficiency levels

The informants' use of songs with different class types

As previously explored, informants' views regarding the usefulness of songs differed when considering different class types: While a large majority of respondents viewed songs as useful when teaching general language courses, agreement rates dropped significantly when considering academic and specific-purpose courses. These views were overall also reflected in the actual use reported by the informants responding to the question: Do you use songs in your classes when teaching the following class types? Table 3.14 illustrates: Over 90% of informants used songs in general second and foreign language courses, and still close to 90% used them in conversation classes. As expected, use for academic and specific purposes was significantly lower. However, a surprising 50% of respondents still used songs in their academic language classes, and over one third in courses for specific purposes. Again, differences in the total percentages compared to previous items can be explained by the fact that only informants who actually taught these levels were considered.

CLASS TYPE	Yes (N)	Yes (%)	No (N)	No (%)	Total applicable
General SL classes	267	96.4	10	3.6	277
General FL classes	261	95.6	12	4.4	273
Conversation	239	88.5	31	11.5	270
Academic purposes	109	50.2	108	49.8	217
Specific purposes	64	38.8	101	61.2	165

Table 3.14: Use of songs with different class types

3.4.2 Teaching goals: For what purposes do informants use songs?

The second research question focused on teachers' goals when using songs and asked: For what purposes – if any – do teachers use songs in the classroom? 374 song-users responded to the questionnaire item exploring this question. Table 3.15 shows how many informants selected the 13 prefabricated answer choices.

What stands out is that many informants use songs to positively influence their students' attitude and learning environment. Over 90% of informants reported that they used songs to motivate their students with an enjoyable activity. One informant elaborated that "music is for the soul" and that "many people thank [him] after such a lesson". Other respondents reported that they used songs to create a sense of variety and change of pace "to break monotony." To create a relaxing atmosphere was the goal of 74.3% of informants. As open-ended answers indicated, a relaxed atmosphere was widely seen as aiding learning success, creativity and interaction. This was summed up by one informant: "Songs allow (...) [students] to relax and this is when they come up with suggestions and ideas which eventually are exchanged with other students."

Response	Ν	%
To motivate students with an enjoyable activity	343	91.7
To practise listening comprehension (and other listening skills)	320	85.6
To create a relaxing atmosphere	278	74.3
To teach authentic language and culture	275	73.5
To introduce new vocabulary	259	69.3
To accommodate different learning styles, e.g. auditive, kin-aesthetic and musical learning styles	257	68.7
To practise familiar vocabulary	231	61.8
To teach multi-word units, i.e. idioms and phrasal language	212	56.7
To teach pronunciation and prosody (including language rhythm)	210	56.1
To introduce new or practise familiar grammatical items	194	51.9
To practise fluency in speaking, i.e. producing spoken language	188	50.3
To give students the opportunity to produce language without feeling observed	171	45.7
As a prompt for spoken interaction, e.g. a class or pair discussion	170	45.5
As a prompt for a writing assignment, e.g. an essay, poem or letter	134	35.8

Table 3.15: Informants' purposes of using songs (prefabricated answer choices)

As expected, a majority of respondents (85.6%) used songs to teach listening comprehension and other listening skills. Another common goal of song-use was the teaching of authentic language and culture. As one informant explained, "[songs] bring authentic language to the class. This is a great chance for students to be away from bookish English". It seems that songs were also utilised by many informants to respond to their students' needs and particular personalities. One frequently selected purpose was the accommodation of different learning styles by using songs, for example to "stimulate learners' artistic skills" or to "get students moving". In addition, giving students, particularly those "who are shy", the opportunity to produce language without feeling observed during the sing-along, was a teaching goal that 45.7% of informants agreed with.

The teaching of vocabulary was also a frequent purpose. 69.3% of informants responded that they used songs to introduce new vocabulary, while 61.8% used them to practise familiar vocabulary. 56.7% saw the teaching of multi-word units as one of their teaching

goals. The fact that between half and more than two-thirds of informants used songs to teach vocabulary in some way warrants the focus of the present research project on the use of songs to foster of word knowledge. In comparison, the number of informants using songs to teach grammar was clearly lower. Respondents also used songs to teach different speaking skills. Teaching pronunciation and prosody was the purpose of 56.1% of respondents, while still half of the informants selected the practicing of fluency as their intended goal. In one informant's opinion, "prosodic features of language are not easily noticed ... Using songs (...) has helped raise (...) awareness to the poetic, rhythmic, metric and intonation patterns". Songs were also used as prompts for other types of activities, for example to initiate spoken interaction or to prompt a writing assignment. Complementary activities used in connection with a song will be further explored in section 3.4.6. Open-ended answers given by 65 informants led to the establishment of 14 additional answer categories, as can be seen in Table 3.16. While these additional categories have little statistical value due to very low numbers of respondents providing answers that can be subsumed under these categories, they offer valuable insight into the informants' reasoning behind the use of songs. It is also possible that some of these categories, had they been offered as prefabricated answer options, would have procured far greater numbers of agreement.

Purpose	N
To introduce, support or recapitulate topics	11
To teach reading and text analytical skills	7
To respond to the learners' needs, interests and/or their existing knowledge	6
To encourage learners' engagement with songs outside the classroom	5
To teach study, thinking & academic skills, e.g. internet search and note-taking	5
To create a sense of variety or a change of pace	4
As background music while students work or take a break	3
To create a connection between teacher and students and/or among students	3
To activate cognitive processes	3
To teach translation skills	2
To make use of and encourage the use of artistic skills	1
To teach behaviours and desired social skills	1
To teach pragmatics	1
To create a sense of belonging to the target culture in the students	1

Table 3.16: Informants' purposes of using songs (open-ended answers)

Songs seemed to be used as a tool to manage the content of a lesson, that is, the topics that were discussed and the activities that were undertaken. Eleven informants explained that they used songs to introduce, support or recapitulate certain topics, for example "as an introduction to a specific topic", by "using songs with references to topics we're studying, such as (...) women's movement" or "to integrate different topics at the end of a unit". In addition, seven respondents used songs to teach reading and text analytical skills. Six teachers also highlighted the fact that they used songs specifically to respond to their students' needs and prior knowledge, for example by responding to learners' song requests and by using the students' favourite music.

3.4.3 Informants' reasons for not using songs

One survey item examined the following research question: If language teachers do not use songs in the classroom – why not? Of the 85 previously identified non-users, 83 responded to this item. A list of seven pre-fabricated reasons was provided to explain why respondents refrained from using songs. In addition, answers given in the "other option" produced eleven further categories. Since some of these additional categories reached higher agreement rates than some of the prefabricated answer options, both are presented together in Table 3.17. Prefabricated answers are labelled (*PF*), and categories derived from open answers are labelled (*OA*). The cut-off point in Table 3.17 regarding the number of respondents was ten, that is, only answers given by ten or more informants are displayed.

Response (PF = pre-fabricated answers; OA = open-answer option)	N (Total: 83)	%
I can't fit songs in with the official curriculum. (PF)	32	38.6
I can't find any suitable songs. (PF)	21	25.3
Using songs means spending too much time on too little learning outcome. (<i>PF</i>)	21	25.3
I tried using songs but my students did not respond well. (Includes: I asked and they said no.) (<i>PF)</i>	17	20.5
In my teaching context, using songs is considered inappropriate for cultural and/or religious and/or political reasons. (OA)	16	19.3
I don't have the necessary technical equipment. (PF)	13	15.7
I don't think songs are useful for language learning in the classroom. (PF)	10	12.0

Table 3.17: Responses to the question – why don't you use songs?

Overall, selection rates for all answer options are low. It seems that none of the available responses resonated with a majority of respondents.

It is important to note that only ten informants identified with the view that songs are of little use as a teaching tool in class. And only two informants added in the open-ended section that they considered songs to be inappropriate when teaching adults. The reasons most frequently selected were based on (1) a conflict with the official curriculum, (2) the inability to find suitable songs, (3) the time spent using songs considering an uncertain learning outcome, (4) the learners' actual or anticipated negative response, (5) the cultural and political context in which the informants find themselves and (6) a lack of technical equipment necessary to obtain or use songs in class. The influence of the cultural and political context on teachers' teaching practice had not been anticipated as an obstacle to their song use during the survey design. A new category had to be established based on similar open-ended responses by a comparatively high number of informants. These respondents reported teaching in a context in which using songs was considered inappropriate by the school administration and/or the students due to the local culture, religion or political system. One informant, for example, described the students' religious objection to the use of songs: "In [this country], many students feel that song is 'haram' (bad) and so if one student (...) does not wish to have songs in the class I am unable to use them." Another described both a cultural and political situation that did not favour the use of songs: "For political reasons, we live in a country (...) where we are closely monitored for whatever we do or say in our classes. Everything has to be in line with the dominant politico-religious ideology. Songs are considered a threat to the native culture[;] even our [own] songs are prohibited for religious reasons. [We] are accused of treason and [freemasonry] if we are caught teaching [English] songs." Only one of the prefabricated answers was chosen by fewer than ten respondents: Four informants selected that the technical equipment was unreliable and prevented them from using songs in class.

3.4.4 Factors influencing the informants' use of songs

The survey also explored what personal, demographic, institutional and societal factors might influence teachers' use of songs (research question 4). I examined the following factors:

- the informants' personal interest in vocal music,
- the informants' musical expertise,
- the country respondents taught in,
- their nationality,
- their gender,
- and the institution they worked at.

In the following, I report on these factors and their possible correlation with the informants' use of songs. Whenever applicable, a Pearson chi-square test of independence was used to test for possible associations between a potentially impacting factor and the informant's teaching practice. More frequently, however, due to the nature of the data, a linear by linear association chi-square test or a simple comparison of descriptive statistics was called for.

3.4.4.1 The importance of music in the informants' private life

As reported earlier, music was seen as a (very) important part of their personal life by 86% of informants. I investigated whether the importance respondents attributed to music in their personal life interacted with their teaching practice. However, a linear-bylinear-association chi-square test showed that there was no significant association between the informants' rating of the importance of music and their use of songs, χ^2 (1, n = 399) = 2.468, p = .116, Cramer's V = .080.

Informants' listening habits

Another way to assess the importance that vocal music played in the informants' private life was to examine their listening habits. That is, I examined how often respondents reported listening to songs. As described above, over 80% of informants did so frequently, that is, often or very often. A Pearson chi-square test for independence showed a significant association between the frequency of listening to vocal music and the use of songs, χ^2 (3, n = 399) = 12.886, p = .005, *Cramer's V* = .180. Table 3.18

illustrates the differences in the listening habits of song-users and non-users: While over 50% of song-users report listening to music very often, only 32% of non-users selected this category. A significantly higher proportion of non-users (28.4%) selected "sometimes".

Frequency of listening to vocal	Informa soi	nts using ngs	Informants not using songs		
music	%	Ν	%	Ν	
Very often	51.3	163	32.1	26	
Often	32.7	104	38.3	31	
Sometimes	14.5	46	28.4	23	
Never	1.6	5	1.2	1	

Table 3.18: Informants' use of songs and frequency of listening to vocal music

When it comes to the role of music in the informants' personal life, we get mixed results: While both song-users and non-users attribute equal importance to music, their reported listening habits differed. It seems that the question of importance did not adequately assess the informants' interest in songs. Examining the respondents' listening habits seems to give a clearer picture of their interest in vocal music and its role in their life. It can be speculated that informants who are highly interested in songs – as reflected by a high reported frequency of listening to them – are also more inclined to try it as a teaching tool in the language classroom.

3.4.4.2 Informants' musical expertise

As mentioned in the literature review in Chapter 2, Edwards (1997) and Pérez Aldeguer and Leganés Lavall (2012) found that teachers' musical expertise impacted on their use of music in class. However, in the present study, training in vocal music did not seem to be a decisive factor: A linear by linear association chi-square test indicated no significant association between the length of time informants participated in a vocal ensemble and the use of songs. Another way of assessing musical expertise was to ask whether informants were able to read sheet music. Again, a linear by linear association chi-square test showed no significant association between being able to read music and the use of songs.

3.4.4.3 Other internal and external factors influencing teaching practices

Section 3.2.4 reported on the respondents' demographic information. In the following I explore whether some of these factors, more specifically the informants' geographical location (country), institution and gender, had an impact on their teaching practices regarding songs.

Informants' geographical location

As we have seen in section 3.4.3, 16 non-users reported teaching in a context that prevented them from using songs for cultural, religious or political reasons. It, therefore, stood to reason that the country the informants were teaching in at the time had a significant impact on their teaching practice involving songs. In other words, it could be assumed that there is a significant association between the respondents' country and their use of songs. The informants' reported locations were categorised by larger regions in order to make a statistical analysis possible. A Pearson chi-square test for independence indicated a significant association between the current country by region and the use of songs, χ^2 (7, n = 400) = 38.419, p = .000, *Cramer's V* = .310. Table 3.19 shows the informants' locations in nine larger regions in the order of highest to lowest percentage of song-users in these regions. A detailed list of all countries contained within each region is also provided.

Table 3.19 confirms that there are great differences in the ratios of song-users and nonusers in the various regions. The percentage of song-users is high with over 80% for informants from the Americas, Australia and New Zealand, Europe and the Balkans and Southeast Asia. There is already a clear drop in song-use in East Asia. And the Middle Eastern and the African region display clearly less frequent song use, as indicated by percentages of song-users below 60%. Admittedly, the total number of informants employed in an African country was small: Only five respondents identified working on the African continent, and only one of them used songs.

Region (Total N of informants)	Informants using songs (%)	Informants using songs (N)	Current country
Middle/South America (12)	91.7	11	Colombia (1), Mexico (1), Uruguay (6), Venezuela (4)
North America (76)	89.5	68	Canada (49), USA (27)
Australia/New Zealand (97)	87.6	85	Australia (16), New Zealand (81)
Europe & Balkan <i>(49)</i>	83.7	41	Belgium (3), Czech Republic (1), Germany (19), Greece (2), Italy (7), Macedonia (1), Poland (1), Spain (10), Ukraine (1), UK (4)
Southeast Asia (34)	82.4	28	Indonesia (3), Malaysia (28), Vietnam (3)
East Asia (72)	73.6	53	China (8), Japan (51), South Korea (9), Taiwan (4)
Middle East <i>(55)</i>	58.2	32	Egypt (2), Iran (3), Jordan (1), Lebanon (1), Oman (4), Qatar (2), Palestine (1), Saudi Arabia (10), Turkey (1), UAE (29), Yemen (1)
Africa (5)	20.0	1	Algeria (2), Comoros (1), Kenya (1), Sudan (1)

Table 3.19: Informants' country (by region) and the use of songs*

* Kazakhstan (1) and India (1) were not considered for this analysis, due to the very low numbers of informants in the respective regions.

These findings reveal that the informants' use of songs is not simply determined by their views regarding the usefulness of a musical approach. Instead, their teaching practice is clearly influenced by their cultural and political environment. A look at the informants' approval of songs as teaching tools in Table 3.20 highlights this influence: Particularly in the Middle Eastern and the African region there is a mismatch of approval and actual use.

Region (Total N of informants)	Informants (%) who view songs positively (agree/strongly agree)	Informants (N) who view songs positively (agree/strongly agree)
Middle/South America (12)	100	12
Africa (5)	100	5
Southeast Asia (34)	94.1	32
Australia/New Zealand (97)	93.8	91
Middle East (55)	89.1	49
North America (76)	84.2	64
Europe & Balkan (49)	83.7	41
East Asia (72)	80.6	58

Table 3.20: Informants' attitude towards song use by region

As Table 3.20 shows, while almost 90% of informants working in Middle Eastern countries regarded songs to be useful in language teaching, only 58% actually used songs in class. Results from respondents employed in African countries show – despite a very low total number – a similar trend. However, these findings also reveal the downside of a statistical analysis that investigates larger regions rather than individual countries. While 17 informants in Middle Eastern countries reported that their society's attitude towards music prevented them from using songs in class, there is still a fairly high number of song-users to be found in this region, implying that their environment has not adopted a similarly negative attitude. A survey assessing the situation in each individual country is called for. However, this is beyond the scope of this research project.

Informants' institution

Another outside factor that might determine a teacher's use of songs is the institution where he or she works. Table 3.21 details the percentage and number of song-users at different types of institutions. It becomes apparent that songs are used by a majority of informants at all types of institutions. Still, the numbers of song-users vary, ranging from 89.5% song-users at primary schools to 75.1% at tertiary institutions and 66.7% at institutions providing vocational training (however, at the latter, the total numbers were very low).

Institution	%	N	Total
Primary school	89.5	34	38
Self-employed	85.7	42	49
Secondary school	84.7	83	98
Continuing education	83.1	49	59
Language school	78.3	65	83
Tertiary education	75.1	142	189
Vocational training	66.7	8	12

Table 3.21: Song use by institution

It does not come as a surprise that most primary- and secondary-school teachers utilise songs, as it seems that the use of music is seen as more acceptable and appropriate when working with children and adolescents. Overall, the differences in song use across the various types of institutions reflect what has been found earlier (as reported in section 3.4.1); that is, the more professional or academic the teaching goals, the less a teacher might be inclined to or allowed to use songs. This is reflected in the relatively lower numbers of song-users at tertiary and vocational institutions.

Gender

Does the informants' gender play a role when it comes to using songs in the language classroom? 405 informants were available for an analysis of a possible association between gender and both the attitude towards songs and their use in the classroom. Table 3.22 shows approval of song-use of male and female teachers separately. Approval was measured by means of agreement with the statement "songs are a useful tool in the language classroom to foster language acquisition". As the table shows, a higher number of female (91.4%) than male (81%) informants takes a positive stance towards songs as teaching tools. And a linear-by-linear chi-square test indicated a significant association between gender and opinion on the usefulness of songs, χ^2 (1, n = 405) = 9.709, p = .002, Cramer's V = .155.

Response	Male Informants (126)		Female In (27	iformants 79)
	%	N	%	Ν
Agree/strongly agree	81.0	102	91.4	255
Neither, nor	15.1	19	7.5	21
Disagree/strongly disagree	4	5	1.1	3

Table 3.22: Informants' attitude towards song use by gender

Differences between genders were even more pronounced, when it comes to the actual use of songs in class: Table 3.23 shows that relatively more female (84.2%) than male (69.8%) informants used songs in class. And a Pearson Chi-square test for independence indicated a significant association between gender and song-use, χ^2 (1, n = 405) = 11.128, p = .001, *phi* = .166.

Response	Male Informants (126)		Female In (22	iformants 79)
	Ν	%	Ν	%
Yes	88	69.8	235	84.2
No	38	30.2	44	15.8

Table 3.23: Informants' song use by gender

However, before drawing any conclusion, it seemed feasible to test for an association between the country the informants were teaching in and the gender ratio in those regions. I hypothesised that in regions where songs are frowned upon by society, the number of female teachers might also be lower than in other regions. Such a correlation would affect the association between gender and song use. And indeed, a Pearson chisquare test for independence indicated a significant association between the informants' gender and the region they teach in, χ^2 (7, n = 400) = 20.869, p = .004, *Cramer's V* = .228. An overview of the gender ratio for the different regions in Table 3.24 reveals that the regions with the least use of songs, that is, Africa, the Middle East and East Asia, have the highest proportion of male relative to female teachers. Such an impact of the region on the gender ratio is likely to have affected the association between gender and song use. And there is, in fact, an overall tendency for regions with a lower proportion of female teachers to display a lower use of songs.

Region	Song-users	Male (%)	Female (%)	Male (N)	Female (N)
Middle/South America	91.7	16.7	83.3	2	10
North America	89.5	21.1	78.9	16	60
Australia/New Zealand	87.6	24.7	75.3	24	73
Europe/Balkan	83.7	30.6	69.4	15	34
Southeast Asia	82.4	23.5	76.5	8	26
East Asia	73.6	48.6	51.4	35	37
Middle East	58.2	41.8	58.2	23	32
Africa	20.0	40	60	2	3

Table 3.24: Gender ratio by region, compared to the use of songs

It seems, then, that one factor behind the association between gender and song use is the region where informants teach. Still, it remains that female teachers also show a significantly higher approval rate of songs as teaching tools. Therefore, gender does, indeed, play a role when it comes to the use of songs in the classroom, with female teachers having a more positive view of music in language teaching. In summary, internal factors such as the informants' gender and interest in music affect their song use for language teaching purposes. However, external factors such as the country and institution where respondents work seem to influence song use to a greater extent and often prevent respondents from using songs despite an overall positive attitude.

3.4.5 A focus on vocabulary – beliefs and teaching practices

The overall focus of this thesis is the use of songs to foster lexical learning. The corpus study discussed in Chapter 4 and the quasi-experimental intervention study presented in Chapter 5 are concerned with the vocabulary found in songs and learners' retention of words and longer connected text when working with songs. As we have seen in Chapter 1, a recurrent claim in the pedagogical literature is that songs are useful to teach vocabulary, including longer multi-word units (e.g. Jolly, 1975; Baechtold & Algier, 1986; Abbott, 2002; Kerekes, 2015). Informal discussions with language teachers, website searches and a review of various ESL/EFL textbooks also indicated that songs are frequently used to teach aspects of word knowledge. It was, therefore, of great relevance to examine whether the survey participants would confirm the assumption that songs are frequently utilised to foster vocabulary acquisition. Two core research questions, thus, specifically addressed aspects of vocabulary teaching: Do teachers use songs to introduce or practise vocabulary and if yes, how? And according to teachers, what characteristics – if any – make songs suitable for vocabulary acquisition?

3.4.5.1 Introducing new vocabulary

Of the 411 informants (both song-users and non-users), 75.2% agreed or strongly agreed with the statement that songs were a useful tool to introduce new vocabulary. As Table 3.25 shows, only 8.5% viewed the use of songs for this purpose negatively.

Response	N	%
Strongly agree	107	26.0
Agree	202	49.2
Neither, nor	67	16.3
Disagree	33	8.0
Strongly disagree	2	0.5

Table 3.25: Informants' opinion about the use of songs to introduce new vocabulary

Naturally, only song-users were asked whether they actually used songs to introduce new vocabulary. Table 3.26 shows that just over two-thirds of 327 informants did, in fact, use songs for this purpose.

Response	Ν	%
Yes	225	68.8
No	102	31.2

Table 3.26: Informants' use of songs to introduce new vocabulary

3.4.5.2 Practicing familiar vocabulary

The response was even more positive when it comes to practicing familiar vocabulary: As can be seen in Table 3.27, of the 408 informants (both users and non-users), 87.3% agreed or strongly agreed that songs were useful to practise familiar vocabulary. And clearly fewer respondents than in the previous item, only 2.8%, (strongly) disagreed.

Response	Ν	%
Strongly agree	130	31.9
Agree	226	55.4
Neither, nor	41	10.1
Disagree	10	2.5
Strongly disagree	1	0.3

Table 3.27: Informants' opinion about the use of songs to practise familiar vocabulary

Again, only song-users (327) were asked regarding their actual teaching practice. As Table 3.28 shows, an evidently higher number of informants reported using songs to practise familiar vocabulary – compared to introducing new vocabulary. This response rate also differs somewhat from a previous item on the purposes of song use in general (section 3.4.2). In fact, the number of respondents selecting the practising of familiar vocabulary as a pedagogical purpose increased from 231 to 249 informants. It remains unclear why an additional 18 informants at this later point reported utilising songs for this purpose. It can be speculated that they considered their own teaching practices more closely as they progressed through the questionnaire.

Table 3.28: Informants' use of songs to practise familiar vocabulary

Response	Ν	%
Yes	249	76.6
No	76	23.4

3.4.5.3 Characteristics that render songs suitable for vocabulary acquisition

The second research question that specifically targeted vocabulary learning examined the following question: According to teachers, what characteristics – if any – make songs

suitable for vocabulary acquisition? To gain insight into why teachers think that songs are suitable for vocabulary learning, this item offered eight prefabricated answer options. Multiple selections were possible. Table 3.29 displays the agreement of 405 informants with these answer options.

Table 3.29: Aspects rendering songs suitable for vocabulary learning (prefabricated answers)

Response	Ν	%
The fact that one can listen to a song several times without getting bored	287	70.9
The fact that songs usually contain authentic language	266	65.7
The rhythm	244	60.3
The possibility to sing along	243	60.0
The fact that one can sing a song several times without getting bored	222	54.8
The rhyme	217	53.6
The fact that learners are likely to encounter a song we used in class outside of class	206	50.9
The melody	198	48.9

An open-ended response option allowed informants to elaborate on their responses and to point out additional factors not considered in the prefabricated answer list. I coded the open-ended answers on the basis of a common reason that, according to the respondents, renders songs suitable for vocabulary learning. These categories are presented below in Table 3.30. The rates of selection were naturally low as no prefabricated choice was provided. The cut-off point was set at a minimum of five respondents who gave an answer that fit the category.

Response	Ν
Songs have a mnemonic effect.	13
Songs are enjoyable and provide a break from the usual routine.	10
Songs contain relevant themes and topics from the target culture (and the related vocabulary).	8
The language in songs is repetitive (words, phrases, chorus).	5
Songs meet existing learner interests.	5

Table 3.30: Aspects rendering songs suitable for vocabulary learning (open answers)

As can be seen in Table 3.29, the three genre-specific features of songs – melody, rhythm, and rhyme – did not elicit the same amount of agreement. Surprisingly, the melody as a relevant factor received the lowest agreement not only compared to rhythm and rhyme but of all prefabricated answer options. A larger number of informants considered the rhyme in the lyrics to be an important feature. The rhythm, however, scored the highest agreement (60.3%). This finding was of particular relevance for the design of the subsequent intervention study. The value that survey respondents attributed to the rhythm highlighted the importance of examining the effect of rhythm without melody. As a result, the intervention study discussed in Chapter 5 not only compared the effect of a sung and a spoken text on verbal learning but also included a poetic spoken format, featuring an explicit rhythm, explicit sound patterns and a poetic written format (see Chapter 5 for the details of the poem format).

I also anticipated that the characteristic "repeatability" of songs might be an important factor. In other words, songs can be listened to and sung several times without becoming boring or tedious. This cannot be said of other text types, even short ones. For example, it does not seem to be common practice to read even a short story more than once or twice. The possibility to listen to a song repeatedly without getting bored received the highest agreement (70.9%) of all prefabricated response options. As one informant put it: "The advantage music has is that it can be listened to over and over again." The productive side of this argument, however, did not quite find the same amount of agreement. Still, over half of the informants found the possibility of repeated oral production to be of particular relevance, and one respondent elaborated that "students are more likely to sing to themselves for pleasure than repeat lists of vocab studied in class". A further five informants provided open-ended answers that stressed the "repetitive nature of songs", that is, they emphasised that the language in songs was itself repetitive due to repeated words, phrases, verses and the chorus.

The fact that songs contain authentic language was an important factor for nearly two thirds of informants. Eight informants also stressed the fact that songs contained relevant themes and topics from the target culture. One respondent emphasised the link between the cultural theme and the vocabulary found in songs, explaining that "the context of songs often makes the link between the vocabulary and cultural aspects more

accessible". The possibility to sing along was also seen as an advantage for vocabulary learning by 60% of informants. Furthermore, 51% of respondents agreed that songs were suitable for vocabulary learning because learners were likely to encounter them again outside the classroom. Unfortunately, the wording of this response is rather imprecise: It remains unclear whether informants see such out-of-class encounters as an advantage in terms of repeated exposure to the lyrics or whether they thought this referred to a consideration of the learners' existing interests. In fact, in the open-ended section five informants described that they used songs for vocabulary learning because songs met their students' existing interests. They reported that the learners were often "curious about what the songs they like are about" and "more determined to learn the language so that they can understand the song". One teacher described how students showed initiative in this respect, as "very often trainees come to me and ask the meaning of the title or a word they heard in a song." Further 13 informants provided open-ended answers regarding their conviction that songs have a particular mnemonic effect. In addition, ten respondents expressed their view that songs aided vocabulary learning because they were enjoyable and provided a break from the usual routine. As these informants explained, the use of songs allowed students to "practice vocabulary in an enjoyable manner", helped them "stay focused for longer periods of time", and sometimes students did not even "realise they [were] learning new words".

3.4.6 How are songs being used by informants?

Research question number 7 addressed the details of song use and asked: How do language teachers use songs in the classroom? The results of seven questionnaire items are reported below. They were considered to be particularly relevant because no previous study had explored the teaching practices involving songs in detail, including the time spent on a song, the typical number of repetitions and popular choices of complementary activities. Again, the findings presented in this section informed the intervention implemented in the quasi-experimental study presented in Chapter 5.

3.4.6.1 Class time spent on a song and related activities

The majority (57.3%) of 328 informants spent no less than 15 minutes on a song and related activities in one lesson. As Table 3.31 shows, most respondents spent 5 to 15

minutes or 15 to 30 minutes. Only a small number of informants spent 5 minutes or less on a song, a timeframe that allows for little more than listening to a song once (see Chapter 4 for the average length of a typical song used for language teaching).

Class time	N	%
5 min or less	20	6.1
5 to 15 min	120	36.6
15 to 30 min	117	35.7
30 to 50 min	58	17.7
over 50 min	13	4.0

Table 3.31: Class-time spent on a song and complementary activities

3.4.6.2 The number of sing-alongs

As we have seen in section 3.4.5, the possibility to sing along is seen as an advantageous characteristic of songs, at least when it comes to vocabulary learning. And as Table 3.32 indicates, the majority of 330 informants (81%) did at least one sing-along in class. The table also details that a small majority opted for one sing-along, while slightly fewer informants did two sing-alongs. Interestingly, almost 20% of informants choose not to sing along at all.

Frequency	N	%
Never	63	19.1
Once	112	33.9
Twice	97	29.4
Three times	40	12.1
More than three times	18	5.5

Table 3.32: The number of sing-alongs

3.4.6.3 The number of repetitions per lesson

Previously, it had been shown that the "repeatability" of songs was ranked as the most important aspect of songs to render them suitable for vocabulary teaching. The results to the present item indicate that informants make, indeed, good use of this possibility to listen to a song several times without getting bored. Table 3.33 provides an overview of how many times informants reported replaying a song in one lesson: A majority of 329 informants played a song three times during one lesson, with the next most popular choice being two repetitions. A surprising 22% of respondents played a song even four or five times.

Frequency	Ν	%
Once	19	5.8
Twice	104	31.6
Three times	134	40.7
Four times	48	14.6
Five times	24	7.3

Table 3.33: The number of times informants repeat a song per lesson

3.4.6.4 The replaying of a song in a subsequent lesson

Distributed rather than concentrated practice is understood to aid long-term retention. In a meta-analysis of 317 experiments, Cepeda, Pashler, Vul, Wixted, and Rohrer (2006), for example, found "that separating learning episodes by a period of at least 1 day, rather than concentrating all learning into one session, is extremely useful for maximizing longterm retention" (p. 370). I was therefore interested whether informants played a song not only once or repeatedly during one teaching unit but also whether they replayed it during a subsequent lesson.

Response	Ν	%
Always	20	6.1
Often	62	18.9
Sometimes	152	46.3
Never	94	28.7

Table 3.34: The replaying of a song in a subsequent lesson

As Table 3.34 shows, only few respondents always repeated a song during a later lesson, and only 19% claimed to replay songs often. The majority conceded that they sometimes return to the same song on another day.

3.4.6.5 Activities used with songs

This item explored the activities most frequently used in direct relation to a song. Eight prefabricated response options were provided. Multiple answers were possible. A total

of 334 informants responded to this item. Table 3.35 gives an overview of the prefabricated answers and the rate of selection by respondents. As can be seen in the table, the most popular activity was the *gap-fill activity*, closely followed by the *sing-along* and the broad category of *discussion*. *Comprehension questions* and *ordering activities* were used by substantially fewer but still over and close to half of the informants respectively. Other proposed activities were used as well, though not as widely.

Activity	N	%
gap-fill activity	249	74.6
discussion	237	71.0
sing-along	235	70.4
comprehension questions	187	56.0
ordering activity	164	49.1
writing activity	131	39.2
true/false statements	101	30.2
dictation	88	26.4

 Table 3.35: Activities used with songs in class (prefabricated answer options)

3.4.6.6 The learners' response to songs

In order to examine how informants experienced their students' response to the use of songs in class, one questionnaire item asked: How would you describe the learners' general response to the use of songs in the classroom? 328 informants observed only a positive, very positive or neutral response, as can be seen in Table 3.36.

Response	N	%
Very positive	141	43.0
Positive	173	52.7
Neutral	14	4.3

Table 3.36: The informants' description of their learners' response

Of course, all these informants used songs and might not have done so if they had experienced more negative learner attitudes. In contrast, 17 informants who reported not using songs argued that students had not responded well (see section 3.4.3).
3.4.6.7 Problems experienced when using songs

I was also interested in the problems that occurred when teachers use songs in class. 328 song-users overall did not report a lot of problems. No prefabricated response reached high numbers of agreement. Admittedly, the answer items were formulated in a fairly strong way. As one respondent correctly pointed out, they should have been more moderate in tone, talking about "some learners ..." or "it is difficult to find suitable songs" rather than "the learners ..." and "I cannot find suitable songs". While this needs to be acknowledged, it seems that the majority of respondents interpreted these answers in less extreme terms. Table 3.37 shows the prefabricated answer choices in order of highest to lowest rate of selection by informants. Multiple selections were possible.

Response	N	%
I cannot find any suitable songs.	94	28.7
The learners do not like to sing.	89	27.1
The noise bothers other classes/teachers.	78	23.8
Too time-consuming.	68	20.7
The learners do not consider songs to be adequate		
for language learning in the classroom.	54	16.5
The use of songs conflicts with the curriculum.	45	13.7
The equipment is unreliable.	43	13.1
We do not have the appropriate equipment.	38	11.6
Problems with learner-discipline.	35	10.7

	• • •	•			
$Iahlp \exists \exists I \cdot Prohlems py$	nerienced wher	n n n n n n n n n n	in class l	nretahricated	answerst
		i using songs	111 61633 [prejubriculeu	unsvers

Against my expectations, only a small number of respondents selected that the use of songs conflicted with the curriculum at their institution. However, in that case it can be assumed that teachers would refrain from using musical activities altogether (see section 3.4.3 on the reasons for not using songs). As can be seen in Table 3.37, the most frequent problem was the difficulty to find suitable songs. Another challenge experienced by a comparatively large number of informants was the high time investment required when preparing and implementing song-based lessons. The fact that 94 and 68 informants respectively experienced these problems lends support to Edwards' (1997) and Perez Aldeguer and Leganés Lavall's (2012) call for the development of more official teaching material involving music and stating clear didactic goals and procedures. Another

frequent problem is related to negative learner reactions as some students did not like to sing or did not consider songs to be *proper* teaching material. This problem might, in fact, also be addressed by means of providing official musical teaching material and by generally fostering an increased acceptance of songs as a teaching tool. However, something else stands out: The finding that 27% and 17% of informants respectively reported having experienced these problems seemingly conflicts with the informants' response to a previous item asking. When asked earlier to describe their learners' general response to song use, none of the informants previously reported a negative reaction. One possible explanation is that these contradictory responses are the result of how the earlier item was worded: Informants were asked to describe "the learners' general response", which did not allow for a more differentiated answer but forced respondents to describe an overall response and to ignore the occasional negative response. Asking the informants about problems they might have experienced when using songs allows for more differentiated reporting.

Other reported problems seem to indicate a lack of adequate facilities and equipment. Informants reported a lack of (reliable) equipment or raised the issue that the noise produced when using songs bothered other classes and teachers. In addition, informants reported on a range of additional problems in the open-answer option. I coded these answers on the basis of a common source of the problem mentioned by the respondents. These problems are presented below in Table 3.38. The rates of selection were naturally low as no prefabricated choice was provided. The cut-off point was set at a minimum of five respondents who gave an answer that fit the category.

Response category	N
I have experienced problems because the society in my country views songs critically due to cultural or political reasons.	11
I have experienced problems due to taste differences between the teacher and the students or among students.	11
I have experienced problems because the school administration, other teachers or parents view the use of songs in class critically.	6

Table 3.38: Problems experienced when using songs in class (open-answer categories)

As Table 3.38 shows, difficulties due to cultural or political reasons are a recurrent theme in this survey. Such difficulties can result in the informants' facing negative reactions from superiors, colleagues, parents or learners when using music in class or, as described in section 3.4.3, they might prevent teachers from using songs altogether. Interestingly, differences in taste also seem to be of importance when it comes to song and music, maybe more so than when it comes to non-musical material. A small number of informants highlighted a "generational gap" or the fear to be "out of sync with [their] students". In summary, the use of songs in the classroom produced fewer difficulties than I expected. While song-users reported a variety of problems, none reached very high rates of agreement.

3.4.7 Song recommendations

In this survey, informants also provided a large number of songs that they used in their classes, as one questionnaire item asked informants for three song recommendations (title and artist) and to specify the purpose of song-use, the reason for recommending these particular songs and the target learners' proficiency level. ESL and EFL teachers in particular recommended a total of 360 different songs. These English songs together with a number of songs gathered in a textbook survey and website search were compiled into a song corpus. An analysis of the pedagogical corpus provided a profile of a typical song used in the English classroom (see Chapter 4). Both the information regarding the typical lesson and the typical song informed the quasi-experimental intervention study which investigated a possible mnemonic effect of songs compared to poems or prose texts (see Chapter 5).

3.5 Discussion

The online questionnaire explored teachers' cognitions regarding the use of songs in the language classroom and their teaching practices involving songs. In other words, it investigated whether second and foreign language teachers used songs and if yes, how and for what purposes. It also examined internal and external factors influencing teachers' choices regarding a musical teaching approach. Finally, this study explored the use of songs particularly to teach vocabulary. It investigated whether informants employed songs to introduce new and practise familiar words. This included the teaching

of multi-word units. Currently, no other questionnaire study has focused on teachers' use of songs to foster lexical learning. Additionally, informants where asked about aspects that, in their eyes, rendered songs suitable for vocabulary learning. A total of 568 teachers working in 41 countries and teaching 20 different languages participated in the study. The wide scope of informants in terms of the languages they taught and the countries they worked in sets this study apart from previous research, which focused on English instructors only and was limited to a small sample of teachers in either the USA (Edwards, 1997; Bjorklund, 2006) or Spain (Pérez Aldeguer and Leganés Lavall, 2012).

The results of the present study revealed seven main findings:

- A majority of informants had a positive view of songs and saw them as useful tools to foster language acquisition.
- 2. A majority of informants used songs in class.
- 3. The average reported teaching unit around a song was fairly elaborate, involving several repetitions of the target song and a variety of complementary activities.
- 4. Informants used songs for a variety of purposes. Their overall goals included creating a positive and motivating learning environment, addressing the learners' individual needs, providing learners with authentic language and culture and teaching clearly defined language skills and linguistic knowledge.
- 5. A majority of informants used songs to introduce new or rehearse familiar vocabulary and to teach multi-word items.
- 6. Informants understood the following characteristics of songs to render them suitable for vocabulary learning: the "repeatability" of song use, that is, the possibility to receptively or productively repeat them several times, their authentic language, their rhythm, the option to orally produce them in the form of a sing-along, the rhyme, their ubiquity outside the classroom, and the melody.
- Non-users, that is, informants who refrained from using songs often did so because of external factors rather than due to a personal conviction that songs are unsuitable as teaching material.

I will now address these findings in greater detail and relate them to the existing literature.

3.5.1 If and how songs are used in language classrooms

The questionnaire results showed that a majority of 447 informants, that is 88%, perceived of songs as a useful or very useful tool to foster second language acquisition inside the classroom. Also, most respondents (391 or 82%) did use songs in the classroom. This is in accordance with Edwards (1997) and Bjorklund (2002), who showed that a majority of their informants, ESL teachers in the USA, utilised songs as teaching material. However, it contrasts with Pérez Aldeguer and Leganés Lavall's (2012) teacher survey, which revealed that the infrequent use of musical activities in class by their informants, EFL teachers in Spain, did not match their overall positive attitude towards songs. It needs to be acknowledged that the data collection and in consequence the results of the present study were potentially skewed towards a positive stance, as proponents of song use were probably more likely to participate in a survey about the use of songs, even though the title "The Use (or Non-Use) of Songs in the Language Classroom" explicitly mentioned non-use in an effort to also solicit responses from teachers with less positive views about the use of songs for language teaching purposes.

The questionnaire data also provided detailed information regarding the teaching practices involving songs. It appears that the typical teaching unit involving songs lasts longer than 15 minutes and involves one to two sing-alongs. A song is typically played three or two times, including the sing-along(s) but only sometimes repeated during a subsequent lesson. Typical activities during the teaching unit are a gap-fill activity, a discussion, a sing-along and comprehension questions. Songs are used with learners of all proficiency levels but most frequently with continuing beginners, low-intermediate learners and high-intermediate learners. In addition, they are frequently used in general language courses such as second and foreign language courses and conversation classes, and to a lesser degree in classes for academic or specific purposes.

As far as I am aware, these findings regarding the details of actual teaching practices involving songs are unique, as no previous study has investigated current procedures involved when teachers use songs in classrooms around the world. However, in many ways these findings correspond with some of the recommendations and personal observations made in the pedagogical literature. Songs are recommended for learners of all proficiency levels (e.g. Arleo, 2000; Abbott, 2002; Lems, 2005; Kerekes, 2015), and the

gap-fill activity, highly popular among the questionnaire informants, has been called "the most overused activity for the detailed task stage of a lesson based on a song" (Hildred, 2011, p. 56; see also Abbott, 2002; Lorenzutti, 2014). It is also frequently proposed that songs are ideal for language learning due to the possibility to repeat them several times and to effectively do "drill exercises" without creating boredom among students (e.g. Richards, 1969; Jolly, 1975; Baechtold & Algier, 1986; Lake, 2002; Lüke, 2008).

3.5.2 Teachers' purposes of song use

The informants in the present study used songs for a variety of purposes which can be separated into four broad categories. They utilised musical activities (1) to create a positive and motivating learning situation, (2) to accommodate individual learners' needs, (3) to provide learners with authentic language and culture, and (4) to teach clearly defined language skills and linguistic knowledge. Music and pop songs are ubiquitous and clearly a highly relevant feature of many societies. It has been argued that particularly for adolescents pop songs are an important means to define and express their identity (North, Hargreaves, & O'Neill, 2000). In addition, several student surveys have indicated that both adolescent and adult learners find the use of songs in class enjoyable, relaxing and motivating and prefer them to other materials (e.g. Jolly, 1975; Gatti-Taylor, 1980; Green, 1993; Bjorklund, 2002; Ludke, 2010). At the same time, it has been found that anxiety specifically in the context of second language acquisition but also the learners' motivation and attitudes towards the learning situation affect learning success (cf. Horwitz, Horwitz, & Cope, 1986; Gardner, Tremblay, & Masgoret, 1997). The present study indicates that teachers use music to foster positive affect, which is believed to be conducive to learning. As the questionnaire revealed, informants often used songs as a motivating, enjoyable and relaxing activity. They also used songs to meet their learners' needs by accommodating different learning styles, but also by providing opportunities for learners to produce language without feeling observed during choral singing. This finding is in accordance with Edwards' (1997) observation that teachers use songs to lower students' emotional inhibitions or, in other words, their affective filter (Krashen, 1983). At the same time informants use songs to provide learners with encounters with authentic language and culture. Finally, respondents used songs to teach

clearly defined language skills and linguistic knowledge, such as listening comprehension, speaking fluency, vocabulary, pronunciation and prosody, and grammar.

3.5.3 Teachers' use of songs to teach vocabulary

Teaching vocabulary was one of the purposes for using songs for a majority of informants, whether in the form of introducing new words or practising familiar words and thus reinforcing and deepening knowledge of the latter. As discussed in Chapter 1, full knowledge of a word comprises various and often incremental steps beyond a first form-meaning connection (Nation, 2013). The recycling of previously introduced vocabulary can, therefore, be considered critical for the consolidation and further expansion of word knowledge (Schmitt, 2008; Nation, 2013). The use of songs to rehearse familiar, that is, at least partially known words, ranked highly among informants. Hence, it appears that teachers utilise songs as a means to reinforce existing word knowledge. In addition, it seems that teachers take advantage of the contextualised presentation of target words in the lyrics to foster the acquisition of "deeper" aspects of word knowledge such as grammatical functions, collocations and associations. This hypothesis also corresponds with the finding that 212 or almost 60% of respondents used songs to teach multi-word units such as idioms and phrasal language.

While no previous teacher survey has focused on vocabulary teaching through songs, it is a frequently proposed purpose of song use in the pedagogical literature (e.g. Jolly, 1975; Baechtold & Algier, 1986; Griffee, 1988; Riddiford, 1998; Abbott, 2002; Esa, 2008). Abbott (2002), for example, claimed that "the repetitive nature of the lyrics can promote the learning of formulaic chunks of language that can be used as readymade expressions in future conversations" (p. 10). When informants were asked about the aspects which, in their view, rendered songs suitable for vocabulary learning, repetition was considered important. Many informants understood it to be beneficial that one can repeatedly listen to and repeatedly sing along with a song without getting bored. In addition, half of the respondents found it advantageous that their students were likely to encounter the songs also outside the classroom. Repetition has been a staple ingredient of learning and memorization since Ebbinghaus' (1885) seminal study on rote learning of word lists. Several studies on the retention of songs in a native language have provided evidence that frequent repetition is also essential when memorizing lyrics (Calvert & Tart, 1993;

Wallace, 1994; McElhinney & Annett, 1996; Moussard et al., 2012). According to a majority of respondents, other aspects that render songs suitable for vocabulary learning were the authentic language found in the lyrics, the rhythm, the possibility to sing along and the rhyme. Surprisingly, only 49% of respondents found the melody to be important. This list of assumed favourable features corresponds largely with findings in cognitive-psychological studies which highlight the positive effect of rhythm, structural regularities such as rhyme, and melody on text retention in the L1 (Chazin & Neuschatz, 1990; Calvert & Tart, 1993; Wallace, 1994; McElhinney & Annett, 1996; Rainey & Larsen, 2002; Tillmann & Dowling, 2007; Purnell-Webb & Speelman, 2008).

3.5.4 Teachers' reasons for not using songs

When it comes to the reasons why informants refrained from using songs, the questionnaire results showed that this was often due to external factors and frequently in spite of a positive attitude towards songs, resulting in what (Borg, 2003) called an "incongruence between cognition and practice" (p. 82). While 85 informants reported not using songs as teaching material, only ten of them declared that they thought songs were not useful for language learning in class. Instead, they reported a variety of other reasons why they would not or could not implement musical activities.

These reasons were in the majority of cases external factors and included the lack of prepared materials corresponding with the official curriculum and a lack of adequate and reliable equipment. In other words, they often could not find any suitable songs and songs that fit in with the official curriculum. Occasionally they also did not have access to the necessary (and reliable) technical equipment. In addition, some reported that the use of songs involved spending too much time on too little learning outcome. This finding supports Edwards (1997) and Pérez Aldeguer and Leganés Lavall (2012), who found that teachers' implementation of a musical teaching approach was strongly affected by the scarcity of official musical teaching material. However, while these studies also found that a lack of official training in the use of music prevented teachers from using songs, the current questionnaire did not find a correlation between the informants' musical expertise and their use of songs. It can be speculated that some non-users would be more inclined to use songs if they had access to ready-made teaching material as part of the official curriculum, including suitable songs and additional activities with clearly

defined teaching goals and pedagogical procedures. Having access to such material would also reduce the time and effort spent on preparing and conducting the lesson.

Another factor preventing informants from using songs was the negative attitude towards music by other "stakeholders", that is, their students, the administration and the society they lived and worked in. More specifically, sixteen informants raised the unanticipated issue that songs were considered inappropriate in their teaching context due to political or cultural reasons. They elaborated that music and songs were deemed offensive based on religious arguments or that foreign songs were viewed as a threat to the native culture. The finding that songs cannot be used in some countries or regions due to political and/or cultural reasons was rendered possible by the international orientation of the present study and is highly relevant in regard to any possible pedagogical recommendations for these parts of the world. None of the previous studies had raised this issue as they were focused on small cohorts of language teachers in Europe and the USA.

In summary, many teachers display a positive attitude towards songs and actually use them as teaching material in class. Rather than viewing them as special treats and time fillers without a specific language-pedagogical purpose, they frequently utilise them with clear goals in mind and in the context of a directed and diverse teaching unit that can include a variety of complimentary activities. One major purpose is the teaching of vocabulary. Of particular interest is the fact that a majority of respondents appears to focus on the consolidation as well as a further deepening of word knowledge, as they frequently utilise songs to practise already familiar words. In addition, the present survey indicated that teachers take advantage of the repeated use of target words in context, reflected by the high use of songs to teach multi-word units.

3.6 Limitations of the survey study

The present survey study has a number of limitations. Most importantly, informants were chosen by means of convenience sampling rather than by a randomised selection process. In addition, as has been mentioned earlier, the survey is likely to have attracted song-users rather than non-users. In other words, even though the title "The Use (or Non-Use) of Songs in the Language Classroom" explicitly mentioned non-use and despite

an effort to also solicit responses from teachers who view the use of songs less positively, proponents of song use were probably more likely to participate in a survey about the use of songs in class. The results of the present study might therefore be skewed towards a positive stance. As a result, it is difficult to generalise some of the findings to the general population of language teachers. Such a skewedness of the data towards the positive also led to difficulties when conducting inferential statistics on some of the questionnaire items. In these cases, the only feasible way of analysing possible associations between items was a comparison of the descriptive data. In addition, the varying completion rates for different items affected the comparability of items as well as the analysis of possible associations between them.

The use of prefabricated answer options also limited the depth of exploring the issue at hand and the extent to which the questionnaire allowed for informants to provide novel and unanticipated information on the topic. However, as the questionnaire was intended to provide a broad understanding of the issue, it needed to strike a balance between the number of questions and the qualitative depth of the questionnaire items and thus the time and effort that respondents were required to invest. Another limitation is the failure to focus on the adult classroom alone. Instead, 38 informants reported teaching at a primary school. Their responses were considered during data analysis.

Furthermore, some of the questionnaire items were of limited value due to difficulties with the wording of the question or the answer options: One questionnaire item used vague prefabricated answers, that is, answers that are highly dependent on each informants' subjective interpretation of the wording (e.g. *very often, often, sometimes*).In addition, the question investigating possible problems occurring when using songs was itself somewhat problematic: Some of its prefabricated answers were too strongly worded, which might have been confusing for some respondents. For example, one answer was: "I cannot find any suitable songs". As informants responding to this item had already identified themselves as song-users, this problem clearly did not deter them from using songs. Hence, a better way of wording this would have been: "I (sometimes) find it difficult to find suitable songs". Finally, respondents pointed out that the survey should have included a section on their learners, as certain learner characteristics such as age and gender strongly influenced the teachers' choices regarding the use of songs.

Overall, this study would have benefited from a separate student survey exploring the topic from the point of view of the learner. However, this was found to be outside the scope of the present thesis, which deliberately focused on teacher cognitions and teaching practices devised and frequently implemented by language instructors.

3.7 Conclusion

This teacher survey explored the use of songs in the language classroom. To this point only few studies had examined this topic, particularly from the teacher's point of view. This survey was, therefore, exploratory in nature, that is, it tried to investigate fairly unchartered territory by answering eight rather broad research questions. Results showed that a large majority of informants had a positive view of songs as teaching tools, and did, in fact, use songs in class. In addition, the survey explored in detail the teaching practices involved when informants implemented song-based teaching units. Typical trends of song-use reported by informants were used to inform the treatment of the intervention study (Chapter 5). Furthermore, results confirmed my hypothesis that many teachers used songs to teach vocabulary, which includes the rehearsal of already semifamiliar words in order to reinforce already existing knowledge and to further deepen word knowledge by presenting words in context, thus affording opportunities to notice and potentially acquire such aspects as syntactical functions and the collocational range of lexical items. This was highly relevant for my thesis as a whole, as the main focus of this research project is the use of songs to foster vocabulary acquisition.

Regarding the lexical profile of song lyrics, informants' voiced a variety of opinions: A majority of informants positively highlighted the use of authentic language in songs. However, there seemed to be somewhat contradictory views regarding the lexical demand of and the vocabulary learning opportunities afforded by lyrics. While most informants supported the use of songs for learners of all proficiency levels, some respondents stated that songs were generally at an advanced level and unsuitable for beginning learners. Others, in contrast, emphasised that song lyrics were generally fairly simple and did not provide much for advanced learners in terms of vocabulary. In response to such a diversity of opinions, I conducted a corpus study and assessed the lexical demand of and the vocabulary learning opportunities afforded by songs used in

the classroom (Chapter 4). The survey also supports a belief widely held by language teachers that songs aid the retention of words, phrases and entire lyrics. One informant asserted that "we all know that songs get stuck in your head." Such a mnemonic effect was frequently attributed to the repetition of (and also in) songs, the rhythm, the rhyme and to the melody. In Chapter 5, I examine such a possible memory effect of songs by means of a quasi-experimental intervention study, comparing learners' verbatim recall of text when working with a song, a poem or a prose text.

Chapter 4 The pedagogical song corpus study

4.0 Introduction

This pedagogical song corpus study explored the lexical demand of songs as well as the vocabulary learning opportunities afforded by songs used in the English classroom. In other words, this study was intended to establish how many words learners need to know to engage with songs and how the lexical profile of songs chosen by teachers or provided in textbooks can best serve lexical learning. As the teacher survey in Chapter 3 revealed, teachers reported using songs to introduce new vocabulary, to practise familiar vocabulary and to teach multi-word items. At the same time informants stated that they experienced difficulties in finding songs that were suitable for their students, including songs containing level-appropriate vocabulary. Informants related conflicting views regarding the lexical profile of authentic songs and described the language of lyrics as too simple or too advanced, too old-fashioned and lacking in everyday or in academic language or in vocabulary for specific purposes. Such a diversity of opinions raises the question how many words learners really need to know to understand authentic songs used as teaching material. It also raises the issue of how the level of vocabulary typically found in songs selected by teachers can best serve lexical learning. Murphey (1989, 1990a, 1992) argued that pop songs were short, lexically simple and highly repetitive. However, his claims regarding the lexical simplicity and repetitiveness of songs were primarily based on the analysis of the type-token ratio of a comparatively small song corpus (13,161 words) and the fact that ten words alone accounted for 25% of tokens in the corpus. While providing some good indication of the results to be expected in the present study, Murphey's findings still require further empirical support. Furthermore, while Murphey assessed the lexical profile of pop songs in general, the present study is concerned with songs used for language-pedagogical purposes.

Finally, teachers can choose from a variety of authentic text genres when selecting teaching material. In order to make an informed decision, particularly when selecting a text to promote lexical learning, it is relevant to understand how the lexical profile of songs compares with those of other text genres. For the purpose of answering these

questions and to establish a lexical profile of songs, a song corpus was compiled and analysed. The corpus also served to provide information regarding a number of characteristics of songs used to teach English. These features potentially influence the language used in the lyrics, the usability of songs as teaching material and the comprehensibility of the lyrics. They included (1) the number of words in a song, (2) the publication date, (3) the English variant spoken by the lead singer(s), (4) the length of the songs and (5) the gender and voice of the lead singer(s). This information regarding characteristics common to songs used in EFL and ESL classrooms was used to pick a "typical" song for the quasi-experimental intervention study discussed in Chapter 5. In this chapter, I will first state the research questions addressed in this study. I will then remind the reader of the methodological approach applied in this study, which is explained in greater detail in the literature review in Chapter 2. I then discuss the compilation and analysis of the present corpus, including the principled selection of songs and the representation of the song lyrics in the corpus. This is followed by a presentation of the results addressing each research question in turn. Finally, I will discuss the study's findings more broadly and point out the limitations of the study as well as opportunities for future research.

4.1. Research questions

The corpus analysis was intended to answer the following research questions:

- 1. What are the characteristics of a typical song used in the English classroom?
- 2. What are the lexical demand of and the vocabulary learning opportunities afforded by authentic, unsimplified English songs chosen by teachers or material designers?
- 3. Do songs intended for use with learners of different proficiency levels differ in their lexical demand?
- 4. How does the vocabulary load of songs differ from that of other authentic text genres used in language teaching?

4.2 Measuring lexical demand and vocabulary learning opportunities

In this song corpus study, I have applied a methodological approach frequently used in a growing number of studies in the field of second language vocabulary research, discussed

in detail in Chapter 2. These studies have investigated how many words learners need to know for adequate comprehension of a genre (e.g. Nation, 2006) and for incidental learning of vocabulary to occur while reading or listening (Webb & Rodgers, 2009a). For written texts, it has been established that the minimal lexical coverage required for reasonable reading comprehension is 95%, while the optimal coverage threshold has been demonstrated to be 98% (Hu & Nation, 2000; Laufer, 2013). No definite threshold has been agreed upon for listening comprehension. However, van Zeeland and Schmitt (2013) and also Bonk (2000) have proposed 95% lexical coverage to be sufficient for comprehension of particular aural text genres, while Stæhr (2009) found 98% to be a reasonable threshold for listening comprehension. For songs, no coverage threshold has been established. As discussed in Chapter 2, songs display characteristics of both written and spoken text and are usually both listened to and read when used as teaching material in the classroom. It, therefore, seems safe to assume that lexical coverage between 95 and 98 percent is a reasonable estimate for comprehension of lyrics.

In order to assess how many words learners need to know to reach this coverage threshold, the present study used Nation's (2004, 2006) word-frequency lists based on the British National Corpus (BNC). Nation's wordlists rank the words found in the BNC according to their frequency, range and dispersion in the corpus and order them by wordlists of 1,000 word families each. The use of Nation's BNC wordlists makes it possible to gauge the frequency of words found in the pedagogical song corpus against the frequency of words in general use of (British) English and to determine the overall vocabulary load of songs selected by teachers and material designers.

4.3 Methodology

This section describes the compilation of the pedagogical song corpus, including the principled selection of songs, the procurement of lyrics and the principled editing process. The corpus is a text corpus and does not include any audio- or video-material. It contains the lyrics of 635 authentic, unsimplified English songs of various genres including pop, rock, rap, country, folk and children's songs. The songs in the corpus were further sub-categorised based on the language proficiency of the intended target learners. 243 songs in the corpus were seen by teachers as appropriate for beginners

(complete and continuing beginners), 356 songs were recommended for intermediate learners (low- and high-intermediate) and only 36 were intended for low- and highadvanced learners. The sub-categorization of songs based on the learners' proficiency levels is further discussed below.

Appendix B provides a list of all songs in the corpus. Due to copyright reasons, the corpus itself cannot be displayed. The lyrics were saved in txt-format. The total number of running words in the complete song corpus is 177,384 – about two-thirds of the size of Webb and Rodgers' (2009b) corpus of TV programmes of 264,384 tokens. A corpus of this size must be considered a small, specialised corpus. However, it is to date the largest pedagogical song corpus compiled for the purpose of establishing a lexical profile of songs used in English teaching. It contains thirteen times more tokens than Murphey's (1990a) song corpus, which comprised 13,161 running words from 50 songs listed in the *Music & Media* Hot 100 Chart from September 1987.

4.3.1. The songs

The songs in the corpus were taken from three types of sources, and their selection followed four principles. The selection process also involved a principled decision on which version or edit of a song to include in the corpus. 635 songs were taken from three kinds of sources:

- 1. The teacher survey (discussed in Chapter 3): 314 songs were recommended by informants responding to the questionnaire;
- 2. ESL textbooks: 62 songs were procured from 26 ESL textbooks (from eleven different series). Only textbooks published in or after the year 2000 were selected to ensure that they were still likely to be used in English classrooms. In addition, the selection was based on availability, that is, on the physical or digital accessibility of the books (or at least their content pages) at libraries or resource archives at Victoria University of Wellington and the Campbell Institute, Wellington, and on websites such as www.slideshare.net and www.amazon.com. For a complete list of the textbooks used see Appendix B.

 ESOL-teacher websites: 257 songs were sourced from six websites, for example www.busyteacher.org, that are used by teachers of English to speakers of other languages (TESOL) to share song recommendations and even detailed lesson plans involving songs. For a complete list of the websites used see Appendix B.

In order for a song to be included in the corpus, four principles had to be adhered to:

- Pedagogical use: The song has been used, is being used or has been recommended for use in the EFL/ESL classroom to teach aspects of the target language.
- 2. Authentic language: The song is authentic, that is, the lyrics have been written for a native-speaker audience (rather than specifically for a non-native-speaker audience).
- 3. Indication of proficiency level: The proficiency level of the target learners has been indicated.
- 4. **Availability**: The lyrics are available in written form, for example on lyrics websites or in CD booklets.

The first selection principle – the requirement of pedagogical use – was considered to be relevant because this study aims to investigate actual classroom practice. Principle 2 required songs to be authentic and unsimplified rather than being written or modified for language teaching purposes. This principle was based on survey findings showing that 74% of survey informants reported using songs to teach authentic language and culture and 66% considered the authentic language of lyrics to be a relevant factor rendering songs suitable for vocabulary learning. In addition, I intended to compare the lexical profile of songs to other authentic text genres used in language teaching, such as novels, newspaper articles, TV programmes and movies.

The third principle – the indication of a proficiency level – turned out to drastically limit the number of sources, particularly the number of suitable ESOL-teacher websites. On such websites, most song recommendations or lesson plans did not include an indication of the proficiency level of the target learners. Instead, in the majority they provided a more precise description of the intended teaching goal, for example the introduction of a

particular grammatical form or specific vocabulary. However, in order to be able to investigate possible differences in lexical demand of songs used with different learners, I considered the requirement of a defined proficiency level to be essential, despite its drawback of limiting the number of suitable sources. While a lexical profile can be established in general terms, the lexical demands on language learners can only be discussed relative to their vocabulary knowledge.

One challenge when using the measure of proficiency level is the actual definition of the different levels. In this thesis – as already described in Chapter 3 – the following terminology was used: absolute beginners, continuing beginners, low-intermediate learners, high-intermediate learners, low-advanced learners, high-advanced learners. These or very similar terms are sometimes referred to as the "classic" terminology of levels description and frequently used to label textbooks. These six stages reflect the six-level-system (A1, A2, B1, B2, C1, C2) of the Common European Framework (CEF), a system of increasing relevance which is "influencing the way in which course book levels are indicated" (Heyworth, 2006, p. 182).

4.3.1.1 The song version selected for the corpus

Many popular songs are published in several versions which differ, for example, in length, exact wording or instrumentation. Such different edits include an album version, a single edit and a radio edit. Fortunately, in many cases, the source indicated a specific version. For example, song recommendations or lesson plans provided on teacher-websites often included the lyrics, a link to a music video or further information about the song. Textbooks also typically displayed the lyrics of a particular edit. However, if no particular edit was specified, the album version was selected.

4.3.2 The procurement and editing of the lyrics

If possible, the song lyrics included in the corpus were taken from CD booklets or the artists' official websites. In addition, some ESOL-websites offered prepared lyrics sheets. Most frequently, however, lyrics were taken from lyrics websites, that is, websites providing the lyrics of a large number of popular songs. The majority of lyrics were obtained from the website *www.lyricsmania.com*. This website was selected for several reasons: An initial comparison of several websites showed that *Lyricsmania* held a

comparatively comprehensive collection of songs from a variety of eras and genres. In addition, a web crawler was used to access and download the lyrics automatically. For this process to be possible, a website needed to meet certain requirements: Firstly, it needed to allow the copying of lyrics. Also, the URLs of the songs provided on a website needed to follow a regular and simple pattern in order for the web crawler to do an automated search based on a list of titles and artists. *Lyricsmania* complied with these requirements. Unfortunately, the lyrics are uploaded by non-professional customers or users on this website. I, therefore, manually checked the corpus for errors. In addition, the song lyrics contained a number of characteristics that needed to be addressed in a principled manner before the analysis. These genre-specific features include frequent repetitions of words, phrases and whole verses, the high frequency of so-called lexical vocables, and the frequent representation of spoken elements in the written lyrics.

4.3.2.1 Choruses, repetitions and fade-outs

The corpus contains all repetitions of individual words, phrases, choruses and verses as used in the selected version of the song – with the exception of fade-outs. An example of word repetition can be found in Amy Winehouse's *Rehab* (Winehouse, 2006): "He's tried to make me go to rehab but I won't go go go." In this case, all repetitions of the word *go* were included. Other repetitions included in the corpus are audible echoes of lines or words vocalised by background singers which do not overlap with the lead-performer's singing. An example can be found in Noa's *Blue touches Blue* (Nini & Dor, 2000): "Blue touches blue (touches blue)." Fade-outs, or a repetition of (parts of) the chorus or verses at the end of the song at decreasing volume, were not included in the lyrics.

4.3.2.2 Non-lexical vocables and marginal words

Song lyrics often contain a high frequency of words such as *oh*, *yeah*, *la*, *ooh*, *deedle*, *doobee*, *shoobee*, *na*. These and similar words are often referred to as non-lexical vocables, as they do not possess lexical content but instead are frequently pronounced for the sake of vocalization itself, to carry the melody and rhythm and to sing without expressing semantic meaning (cf. Chambers, 1980). Despite the occasional resemblance with marginal words, non-lexical vocables do not convey the same referential, social or even emotional meaning as their possible homonyms. (A complete list of all non-lexical

vocables in the corpus can be found in Appendix B.) Non-lexical vocables in the corpus accounted for 5,977 tokens. Given their high frequency, it can be expected that they greatly influence the lexical profile of the corpus. However, they have little or no referential or social meaning and can, thus, be assumed to pose only a small or no learning burden. It was, therefore, decided to exclude them from the analysis. Clearly identifiable marginal words such as *shh*, *oops*, *tada* or *wow*, on the other hand, remained part of the analysed corpus. Non-lexical vocables have an effect on the speed of speech production. They might provide listeners with short mental breaks and might allow for them to reflect on the input, for example by processing previously heard or anticipating upcoming words.

4.3.2.3 The representation of spoken language in the written lyrics

Since written lyrics are often intended to reflect characteristics of spoken language, they often contain a high number of contractions (e.g. *I've, can't*), connected speech (e.g. *shoulda, woulda, coulda*) and apostrophised abbreviations (e.g. *lovin'*). The RANGE programme (Heatley, Nation, & Coxhead, 2002), a computer programme used for the corpus analysis, automatically separates contractions. For example, *can't* is counted as two words, that is, as *can* and *not*. Connected speech, on the other hand, is counted as one word. For example *coulda* is counted as a family member of *can*. Apostrophised abbreviations were manually added to the frequency lists used in the analysis. For example, *lovin'* was subsumed under the lemma *love*. Quasi-transcriptions of spoken variations of words such as *ya* (*you*), *cos* or *cuz* (*because*) where added to their respective word families in the frequency lists.

4.4 Data analysis

IBM SPSS Statistics 19 was used to investigate characteristics of songs chosen by English teachers or material designers. Furthermore, WordSmith Tools (Scott, 2008) was utilized to determine the standardised type-token ratio (STTR) of teacher selected songs. The STTR served to measure the overall lexical diversity of the lyrics. Finally, the computer programme used to analyse the lexical profile of the song corpus was RANGE (Heatley et al., 2002), available on Paul Nation's website:

http://www.victoria.ac.nz/lals/about/staff/paul-nation. The RANGE programme counts

the number of times a word occurs in the corpus. Results are presented as tokens, types and word families. In addition, RANGE lists the word families according to their frequency in wordlists used during the analysis. For this purpose, I used Nation's (2004, 2006) twenty BNC wordlists, which enabled the assessment of the overall vocabulary load of the song corpus and a comparison with other studies utilizing the same methodological approach. For a detailed discussion of Nation's BNC frequency lists, see Chapter 2.

4.5 Results

The present song corpus was analysed in two ways. Firstly, it was examined with regard to characteristics typically found in songs used as teaching material. These characteristics included the average number of words, the typical publication date, the singer's country of origin, the average length of a song and the lead-singer's or lead-singers' voice(s). Secondly, the corpus was analysed in regard to its lexical profile. More specifically, the standardised type-token ratio of the song corpus was established in order to determine the overall lexical diversity of songs used in English classrooms. Furthermore, I assessed the overall vocabulary load of teacher-selected songs. In the following, I will present the results of this threefold analysis and address the four research questions stated in section 4.1.

4.5.1 Common characteristics of songs used in ESL/EFL classrooms

The first research question addressed in this study was: What are the characteristics of a typical song used in the English classroom? It was assumed that certain features might reflect teachers' pedagogical reasoning when selecting songs for in-class use, for example in regard to the number of words and the overall length of a song. It was also hypothesised that certain aspects of songs, such as the singer's gender and the variety of English used, could potentially influence the comprehensibility and the vocabulary of the lyrics. In addition, features found to be typical of a teacher-selected song were used to inform the intervention study discussed in Chapter 5.

4.5.1.1 The average number of words

Song lyrics are a comparatively short text genre: As Table 4.1 illustrates, the average number of words in a song in the present corpus is 289 (median = 269).

	Tokens
	(N)
Average	289
Standard deviation	120
The longest song	854
The shortest song	32

Table 4.1: The average number of words in the corpus songs

The longest song, Justin Timberlake's *What goes around* (Timberlake, Mosley, & Hills, 2006), contains 854 words while the shortest song, the traditional children's song *Head*, *shoulders, knees and toes* only contains 32 words. In contrast, the shortest text in Webb and Rodger's (2009b) corpus of TV programmes was a news story and comprised 523 words.

4.5.1.2 The date of publication

For in-class use, teachers and material designers seem to clearly favour songs that have been published fairly recently. As Table 4.2 shows, 92% of songs were published after 1959 and over one third of the songs were published in or after the year 2000 (36.4%). The most popular decade was clearly the time from 2000 to 2009 with 30.6% of songs published during this period. The 1990s were the second most popular decade with 17% of songs being published during that time. Songs from the 1960s (13.7%) and 1970s (13.2%) were slightly more frequent in the corpus than songs from the 1980s (11.7%).

Period of	Songs in the corpus		
publication	Ν	%	
before 1900	10	1.6	
1900-1909	0	0	
1910-1919	2	0.3	
1920-1929	2	0.3	
1930-1939	5	0.8	
1940-1949	5	0.8	
1950-1959	19	3.0	
1960-1969	87	13.7	
1970-1979	84	13.2	
1980-1989	74	11.7	
1990-1999	108	17.0	
2000-2009	194	30.6	
2010-2011	37	5.8	
no date	8	1.3	

Table 4.2: The period of publication of the songs found in the corpus

4.5.1.3 The singers' country of origin

The variant of English used in a song affects the choice of vocabulary used in the lyrics. To assess the English used in a song, I analysed the country of origin of the singer(s). Table 4.3 shows how many songs in the corpus were performed by bands from North America, that is the USA or Canada (for example Bryan Adams, Elvis Presley and Jason Mraz), from Great Britain (for example the Beatles, Robbie Williams and Amy Winehouse), from Australia or New Zealand (for example Bic Runga, Savage Garden and Shihad), from Jamaica (Bob Marley and Jimmy Cliff) or from South Africa (Zain Bhikha). In addition, a number of songs were performed by singers from countries where English is not an L1. Examples of such non-native singers are ABBA (Sweden), Noa (Israel), Tokio Hotel (Germany) and X Japan (Japan).

Country	Ν	%
USA/Canada	377	59.4
Great Britain	178	28.0
Non-native singers	45	7.1
Australia/New		
Zealand	27	4.3
Jamaica	4	0.6
South Africa	1	0.2
Unknown origin	3	0.5

Table 4.3: The country of origin of the singers

By far the most songs in the corpus are performed by singers from the USA or Canada (see Table 4.3). The second most frequent category – Great Britain – comprises considerably fewer songs. All other categories together (including songs of unknown origin) comprise 12.7% of songs in the corpus.

4.5.1.4 The average length of corpus songs

The average length of the songs in the corpus was just under four minutes (3:50) with a standard deviation of one minute. The mode was 3:40. The longest song, Led Zeppelin's *Stairway to heaven* (Page & Plant, 1971), had a playing time of 8:02 minutes. The shortest song was the *Addams Family Theme Song* (Mizzy, 2010) with only 0:54 minutes.

4.5.1.5 The voice of the singer(s)

The voice of the singer(s) might affect the learners' attention to and recall of orally presented material (e.g. Gruber & Gaebelein, 1979). Particularly the singer's or singers' sex might be considered a relevant variable. Markham (1988), for example, found that ESL learners participating in a study on listening comprehension had a gender bias and "listened more attentively to the male speaker" (p. 397), resulting in higher recall of the content of the target texts. Whether this finding can be extended to popular music is unclear. This study analysed the singer's or singers' sex, using the categories male, female and mixed. Mixed refers to bands with several lead-singers of both sexes. Another category comprised voices of children or adolescents as well as songs with electronically distorted vocals. A majority of songs in the corpus, 362 or 57%, was sung by one or several male lead-singers. Female singers vocalised 206 or 32.4% of the songs in the corpus. Songs with mixed vocals made up only 6.6% of the corpus, that is, 42 songs. Only four songs (0.6%) were sung by children or by a digitally distorted voice, among them Justin Bieber's Baby (Bieber, Stewart, Nash, Bridges, & Milian, 2010) and Eiffel 65's Blue (da ba dee) (Jey, Lobina, & Gabutti, 1998). For 21 songs (3.3%) no specific voice could be determined because the song-source did not include information on a specific performer. This "unknown" category includes such songs as the anthem God defend New Zealand and the traditional folk song Old MacDonald had a farm.

4.5.2 The lexical profile of songs

The second research question that this study was designed to address was: What are the lexical demand of and the vocabulary learning opportunities afforded by authentic, unsimplified English songs chosen by teachers or material designers? In order to respond to this question, I analysed the song corpus using two different tools and established a lexical profile of teacher-selected songs. Firstly, I assessed the lexical diversity of the song corpus by determining the standardised type-token ratio (STTR). Secondly, I determined the vocabulary load of the song corpus.

4.5.2.1 The standardised type-token ratio of the song corpus

A corpus analysis using the RANGE programme (Heatley et al., 2002) revealed that the song corpus comprised 177,384 tokens and 6,496 types. A further lexical analysis of the

corpus using WordSmith Tools, version 5 (Scott, 2008) showed that the corpus has a standardised type-token ratio (STTR) of 27.79, which indicates that it is lexically less diverse and clearly more repetitive than for example the written part of the International Corpus of (British) English (ICE-GB) with an STTR of 38.68, and even the spoken part of the ICE-GB, which has an STTR of 35.85 (cf. Kreyer & Mukherjee, 2007). (The ICE-GB was selected for a comparison because it had already previously been used by Kreyer and Mukherjee (2007) to compare the lexical diversity of a song corpus with spoken and written English.) What is even more interesting is that the vocabulary in the pedagogical song corpus is also more repetitive than the language in Kreyer and Mukherjee's (2007) Giessen-Bonn Corpus of Popular Music (GBoP), which has an STTR of 38.44. We recall from the literature review in Chapter 2 that the GBoP can be described as a corpus representative of pop music in general rather than of songs selected on the basis of a pedagogical rationale.

4.5.2.2 Determining the vocabulary load of the song corpus using the RANGE programme

Another way to measure both the lexical demand of songs and the vocabulary learning opportunities they afford is an analysis by means of the RANGE programme (Heatley et al., 2002) and Nation's BNC wordlists (Nation, 2004, 2006). This methodological approach is described in detail in the literature review in Chapter 2, and is also briefly summarised above in section 4.2. We recall that the RANGE programme counts the number of times a word occurs in the corpus. In addition, the programme lists word families encountered in the corpus according to their frequency in wordlists used during the analysis. The present study gauged word frequencies in the song corpus against Nation's (2004, 2006) 20 BNC wordlists. These wordlists rank words in the British National Corpus (BNC) according to their frequency, range and dispersion in the corpus and order them by wordlists of 1,000 word families each. The use of these wordlists makes it possible to compare the lexical profile of teacher-selected songs to word frequencies in general language use (represented by the BNC) and to measure how many words a learner needs to know to reach a certain lexical coverage of a corpus or text. As discussed in Chapter 2, 95% to 98% was considered to be a reasonable coverage threshold for unassisted comprehension of pop songs. However, it needs to be acknowledged that this lexical benchmark is based on

research investigating reading and listening comprehension of various written and spoken types of discourse but that no coverage threshold particularly for songs has been established.

The RANGE analysis revealed that the 635 songs in the song corpus used 177,384 tokens (excluding the 5,977 non-lexical vocables), 6,496 types and 4,018 word families from the BNC frequency lists. Table 4.4 provides the raw numbers and percentages of tokens, types and word families for each of the twenty frequency levels as well as proper nouns, transparent compounds, marginal words and words not found in any of the lists.

WORD LIST	Tokens		Types	Families	
	Ν	%	Ν	%	Ν
1,000	159,169	89.73	2,181	33.57	882
2,000	9,017	5.08	1,209	18.61	685
3,000	3,136	1.77	800	12.32	532
4,000	1,570	0.89	508	7.82	388
5,000	930	0.52	337	5.19	281
6,000	585	0.33	206	3.17	174
7,000	290	0.16	156	2.40	145
8,000	161	0.09	107	1.65	98
9,000	255	0.14	106	1.63	97
10,000	186	0.10	78	1.20	73
11,000	131	0.07	70	1.08	65
12,000	75	0.04	48	0.74	46
13,000	74	0.04	37	0.57	36
14,000	22	0.01	15	0.23	15
15,000	62	0.03	11	0.17	10
16,000	61	0.03	25	0.38	23
17,000	28	0.02	13	0.20	13
18,000	21	0.01	9	0.14	9
19,000	20	0.01	14	0.22	14
20,000	38	0.02	15	0.23	14
Proper nouns	1,094	0.62	356	5.48	356
Transp. compounds	103	0.06	43	0.66	43
Marginal words	40	0.02	8	0.12	8
Not in the list	316	0.18	144	2.22	?
Total	177,384		6,496		4,018

Table 4.4: The frequencies of words in the song corpus in tokens, types and families

The first 1,000 most frequent word families make up 159,169 tokens or 89.7% of words in the corpus. The second set of 1,000 word families accounts for 5.1%, and the third

most frequent 1,000 word families make up 1.8% of tokens in the song corpus. All wordlists beyond the 3,000 most frequent word families account each for less than 1% of the tokens in the corpus, with a rapid decline over the mid-frequency levels to coverage of less than 0.2% beyond the 7th word frequency band. Proper nouns account for 0.62% of running words. A separate RANGE analysis assessing the number of academic words in songs revealed that words from the Academic Word List (AWL) (Coxhead, 2000) account for 0.33% of tokens in the song corpus.

The second research question was concerned with both the lexical demand and lexical learning opportunities of songs. This question was approached by establishing the lexical coverage necessary to reach adequate comprehension of song lyrics. Table 4.5 shows the cumulative coverage of the twenty BNC frequency lists – with and without proper nouns, transparent compounds and marginal words – of the tokens in the song corpus.

WORDLIST	Coverage	Coverage
WORDEIST	without PN, TC, MW	with PN, TC, MW
1,000	89.73	90.43
2,000	94.81	95.51
3,000	96.58	97.28
4,000	97.47	98.17
5,000	97.99	98.69
6,000	98.32	99.02
7,000	98.48	99.18
8,000	98.57	99.27
9,000	98.71	99.41
10,000	98.81	99.51
11,000	98.88	99.58
12,000	98.92	99.62
13,000	98.96	99.66
14,000	98.97	99.67
15,000	99.00	99.7
16,000	99.03	99.73
17,000	99.05	99.75
18,000	99.06	99.76
19,000	99.07	99.77
20,000	99.09	99.79
Proper nouns	0.62	
Transp. compounds	0.06	
Marginal words	0.02	

Table 4.5: Cumulative coverage, without and with proper nouns (PN), transparent compounds (TC) and marginal words (MW)

As discussed in Chapter 2, proper nouns (PN), such as Rudolph, Billy Jean or Hollywood, and marginal words (MW), such as oops and wow, can be considered known due to their small learning burden (Hwang & Nation, 1989; Hirsh & Nation, 1992; Nation, 2006). Transparent compounds (TC), such as *teardrop*, *backseat* and *nosebleed*, can be easily understood when knowing their high-frequency parts (Bauer & Nation, 1993). Proper nouns, marginal words and transparent compounds can therefore be included when assessing the vocabulary knowledge required to reach certain coverage thresholds. Table 4.5 shows that knowledge of 2,000 word families plus proper nouns, transparent compounds and marginal words is necessary to reach 95% coverage of the song corpus, while 4,000 word families plus proper nouns, transparent compounds and marginal words need to be known to reach 98% coverage. The lexical demand is slightly higher, if proper nouns (and transparent compounds and marginal words) are not assumed to be known. That is, without proper nouns knowledge of 3,000 and 5,000 word families is required for 95% and 98% coverage of tokens in the corpus. When we consider these results in terms of vocabulary learning opportunities rather than lexical demand, we can see that teacher-selected songs contain only few words beyond the 4,000 most frequent words. To put it differently, knowing the most frequent 4,000 words (and proper nouns, marginal words and transparent compounds) provides coverage of 98% of teacherselected songs, which means that in a song of average length, (289 words) only five to six tokens are unknown. This finding has implications particularly for advanced learners, which will be discussed in detail in the discussion below and in the final discussion in Chapter 6.

4.5.2.3 Assessment of the variation of required vocabulary knowledge among songs Given the high number of very short texts in the corpus, the question remains whether the required vocabulary knowledge of 4,000 words to reach 98% coverage is a good representation of the lexical demand of the majority of texts in the corpus. In order to ascertain that the overall lexical demand of 4,000 word families was indeed representative, I calculated the 95%-confidence interval for a subset of 108 songs. While this approach seems somewhat unorthodox – as the continuous data is treated as categorical data - this method can provide an indication whether the average lexical demand of 4,000 word families might indeed pertain to the majority of songs in the

corpus. A subset of 108 was considered large enough to allow a generalization to the complete corpus. The 108 songs consisted of 36 songs from each subcorpus, that is, the beginner subcorpus, the intermediate subcorpus and the advanced subcorpus. While songs from the beginner and intermediate subcorpora were selected randomly, using a random number generator available online at www.random.org, the advanced subcorpus only comprised 36 songs in total and a random selection was not possible. In order to calculate the 95%-confidence interval, the BNC-frequency bands were understood to be categorical data ordered from 1 to 20. Figure 4.1 shows the distribution of the 108 songs across the twenty frequency levels.

Figure 4.1: Word knowledge (in 1,000-word frequency bands) required to reach 98% coverage of 108 songs



The graph in Figure 4.1 illustrates that most songs require vocabulary knowledge between 2,000 and 4,000 word families for coverage of 98% of words in their lyrics. In fact, the statistical average is 4.27. The standard deviation is 3.14, and the measurement error is 0.592274. The lower bound of the 95% confidence interval is 3.676245, whereas the upper bound is 4.860793, indicating that we can be 95% confident that the mean is within the range of 3.7 and 4.9. A required vocabulary size of 4,000 word families is, thus, a good estimate of the lexical demand of the 108 songs and also of the entire song corpus.

4.5.3 The vocabulary load of songs used with learners of different proficiency levels

The third research question was concerned with possible differences regarding the vocabulary load in songs teachers use with learners of different proficiency levels. It asked: Do songs used with learners of different proficiency levels differ in their lexical demand? As discussed in the methodology section 4.3, songs in the corpus were sub-categorised according to the intended proficiency level of the potential target audience. In total, 243 songs in the corpus were intended for beginners (complete and continuing beginners), 356 songs were intended for intermediate learners (low- and high-intermediate), and only 36 were intended for (low- and high-) advanced learners. Table 4.6 shows the total numbers of tokens in each subcorpus, as well as the raw numbers and percentages of tokens accounted for by 20 frequency-lists, by proper nouns, transparent compounds and marginal words. Results for lists 10 to 20 are given as one accumulated number.

WORD LIST	Beginner corpus		Intermediate corpus		Advanced corpus	
	N	%	N	%	N	%
1,000	57,124	90.45	91,655	89.53	10,390	87.69
2,000	3,093	4.90	5,154	5.03	770	6.50
3,000	955	1.51	1,957	1.91	224	1.89
4,000	558	0.88	873	0.85	139	1.17
5,000	296	0.47	576	0.56	58	0.49
6,000	188	0.30	344	0.34	53	0.45
7,000	53	0.08	218	0.21	19	0.16
8,000	40	0.06	106	0.10	15	0.13
9,000	87	0.14	139	0.14	29	0.24
10,000-20,000	237	0.37	437	0.44	44	0.38
Proper nouns	390	0.62	640	0.63	64	0.54
Tr. compounds	22	0.03	17	0.02	1	0.01
Marg. words	35	0.06	52	0.05	16	0.14
Not in the list	80	0.13	<u>2</u> 10	0.21	26	0.22
Total	63,158		102,378		11,848	

Table 4.6: Coverage in the three subcorpora (tokens only)

Table 4.6 shows that the three subcorpora are fairly similar regarding the percentage of proper nouns and the percentage of low-frequency words. Coverage provided by the first 1,000 word families, however, differs between the three subcorpora, with the beginner subcorpus displaying the highest coverage and the advanced subcorpus displaying the lowest coverage, resulting in a difference of 2.8%. The second 1,000 word families, however, show a reversed pattern, that is, they account for the highest percentage of

tokens in the advanced subcorpus (6.5%) and the lowest percentage in the beginner subcorpus (4.9%). This difference might be explained by the comparatively small size of the advanced subcorpus with only 36 songs and a total number of only 11,848 running words. In such a small corpus, high use of only a few words from a specific frequency level can have a great impact on the coverage provided by that wordlist. A look at the most frequent words from the second wordlist used in the advanced subcorpus reveals a high use of words in only one or only few songs: The word *medicine*, for example, is used 23 times in only one song (Bon Jovi's *Bad Medicine*). The word *guilty* is used 21 times in only two songs. However, it is only used once outside of Bananarama's *Love in the First Degree*. Other examples of frequent words with a limited range from the second wordlist are *faith* (18 uses in one song), *strange* (18 uses in two songs), *strength* (18 uses in one song), *brick* (14 uses in one song), *mad* (14 uses in two songs), and *junction* (12 uses in one song).

The three sub-corpora also display differences in the coverage provided by midfrequency vocabulary, that is, words from the 4th to 9th BNC-frequency bands, with the strongest disparity again found between the beginner subcorpus (1.9% coverage) and the advanced subcorpus (2.6%). Do these disparities result in a different lexical demand to reach 98% coverage? Table 4.7 compares the coverage provided by the first 5,000 word families in the three subcorpora and the complete corpus comprising all three levels.

WORDLIST	Complete corpus	Beginner	Intermediate	Advanced	_
1,000 + PN, TC, MW	90.43	91.16	90.23	88.38	
2,000 + PN, TC, MW	95.51	96.06	95.26	94.88	
3,000 + PN, TC, MW	97.28	97.57	97.17	96.77	
4,000 + PN, TC, MW	98.17	98.45	98.02	97.94	
5,000 + PN, TC, MW	98.69	98.92	98.58	98.43	

Table 4.7: Cumulative coverage for the complete corpus and the subcorpora includingproper nouns (PN), transparent compounds (TC) and marginal words (MW)

As Table 4.7 shows, a vocabulary size of 4,000 word families plus proper nouns, transparent compounds and marginal words is required to gain 98% coverage of all three subcorpora. However, the table also illustrates that the three subcorpora can be ranked in terms of lexical demand, with the advanced subcorpus displaying the highest and the beginner subcorpus displaying the lowest demand. Still, 4,000 word families is a surprisingly high coverage for songs that teachers intend for use with beginning learners. Again, a closer look at the range of words indicates that the high repetition of words in only one or very few texts might distort the overall lexical demand of songs in the subcorpus. An analysis using WordSmith Tools (Scott, 2008) reveals that 28 types in the beginner subcorpus appear in only one song but are repeated between ten and 25 times within one song. These items account for 0.6% of tokens. The proper noun *Barbie*, for example, occurs 25 times in only one song, Aqua's *Barbie Girl* while the word *submarine* also occurs 25 times in one song alone, that is, *Yellow Submarine* by the Beatles. Other examples are *liked*, *shorts*, *weekend*, *partyin'*, *Fernando*, *glass*, *replay*, *rocky*, *seasons*, *Colorado*, *ladies*, *MacDonald*, *clink*, *February*, *glug*, *January*, *shh*, *swish*, *jumping*, *monkeys*, *believer*, *compares*, *country*, *deer*, *sandman*, and *technology*.

4.5.4 Comparing the vocabulary load of songs with the lexical demand of other authentic genres

The fourth research question was concerned with the vocabulary load of songs in comparison to other types of discourse used in the English classroom. It asked: How does the vocabulary load of songs differ from that of other authentic text genres used in language teaching? Table 4.8 compares word knowledge necessary to reach 95% and 98% coverage of different genres. These different types of written and spoken discourse have been discussed at greater length in Chapter 2.

As Table 4.8 illustrates, teacher-selected songs require knowledge of comparatively fewer words than other written and spoken genres to reach 95% and 98% coverage. The difference is most marked – up to 4,000 word families – when comparing songs and written texts. Songs and unscripted as well as scripted spoken language, on the other hand, differ by 1,000 to 3,000 words.

Table 4.8: Lexical knowledge required for 95% and 98% coverage of different authentictext genres (PN = proper nouns, MW = marginal words, TC = transparent compounds)

Text Genre	95%	98%	Study
Teacher-selected songs	elected 2,000 + PN, MW, 4,000 + PN, MW, TC TC		the present thesis
Novels (adult)	4-5,000 + PN	8-9,000 + PN	Nation 2006
Novels (adolescent)		5,000	Hirsh & Nation, 1992
Newspaper articles	paper articles 4,000 + PN 8-9,000 +		Nation, 2006
Spoken discourse	3,000 + PN	7,000 + PN	Nation 2006
Spoken discourse	3,000		Adolphs & Schmitt, 2003
Movies	3,000 + PN, MW	6,000 + PN, MW	Webb & Rodgers, 2009b
TV programmes	3,000 + PN, MW	7,000 + PN, MW	Webb & Rodgers, 2009a

However, as discussed in the literature review in Chapter 2, one should also look at the details of lexical coverage provided by the different frequency bands. Table 4.9 shows the coverage provided by high- and mid-frequency words in the different written and spoken text genres.

Table 4.9: Coverage (in percent) provided by high- to mid-frequency words in differentauthentic text genres

	Teacher- selected	Movies	τν	Unscripted spoken	Newspaper articles	Novels
	songs	(Webb & Rodgers, 2009a)	(Webb & Rodgers, 2009b)	(Nation, 2006)	(Nation, 2006)	(Nation, 2006)
1,000	89.73	86.52	85.11			
2,000	94.81	90.67	89.53	89.35	83.00	87.83
2,000 +PN/MW	95.45	94.04	93.52	90.38	87.55- 89.12	
3,000 +PN/MW	97.22	95.76	95.45	96.03		
4,000 +PN/MW	98.11	97.04	96.75		95	94.8
Proper nouns	0.62	2.67	2.96	1.03	4.55-6.12	1.53
Marginal words	0.02	0.70	1.03			

As Table 4.9 illustrates, the first and the second 1,000 word families provide a considerably higher coverage of tokens in song lyrics than in other text genres. (Unfortunately, studies on other genres did not always provide exact coverage data for all frequency levels.) Coverage provided by 2,000 word families differs by between 4.1%, when comparing songs and movies, and 11.8%, when comparing lyrics with newspaper articles. In contrast, the number of proper nouns is clearly lower in the song corpus than in the other written or spoken corpora. The low number of proper nouns is in accordance with Murphey's (1990) finding that song lyrics tend to be deliberately vague and frequently do not provide specific times, places and personal referents. In addition, the number of marginal words is very low compared to other scripted spoken language in TV programmes and movies. This provides some indication that songs are, at least in this respect, distinct from scripted and unscripted spoken discourse. This finding contradicts Murphey's (1990) claim that songs can be considered situational spoken discourse. However, it provides further support for Kreyer and Mukherjee's (2007) proposition that song lyrics resemble written rather than spoken text in certain respects. Kreyer and Mukherjee found that song lyrics, while resembling spoken language in many ways, display a much lower use of certain fillers relevant in interactive and spontaneous spoken discourse. While Kreyer and Mukherjee only analysed the use of the discourse marker you know, the present study shows that their claims also pertain to other exclamations, interjections and hesitation markers, which are clearly less frequent in song lyrics than in TV and movie scripts. Overall, when proper nouns and marginal words are considered to be known, the difference in coverage between songs and other genres is still apparent but less striking.

4.6 Discussion

The present analysis of a pedagogical song corpus investigated the vocabulary load of teacher-selected songs used in EFL and ESL classes with a specific focus on their lexical demand as well as the vocabulary opportunities they afford. The vocabulary load of song lyrics was compared to the lexical demand of other written and spoken types of discourse. Furthermore, this study examined whether the vocabulary load of the song lyrics differed relative to the proficiency levels of the target learners. Finally, it also investigated several characteristics of songs used for language-pedagogical purposes. The

only other existing song corpus analysis with a pedagogical focus is Murphey (1990a, also 1989, 1992). Murphey argued that pop song lyrics are simple, repetitive, short and also vague in terms of times, places and personal referents referred to in the lyrics. Results of the present study show that Murphey's findings also largely prove true for pop songs specifically selected by teachers for use in the language classroom. Regarding the lexical demand, results of the current corpus analysis revealed that songs selected by teachers and material designers for use in class require knowledge of the 2,000 most frequent word families plus proper nouns and marginal words for 95% coverage and knowledge of 4,000 word families for 98% coverage.

These findings indicate that teachers select song lyrics that are clearly less demanding in terms of vocabulary load than other authentic text genres used in the language classroom. We recall that scripted spoken discourse found in TV programmes (Webb & Rodgers, 2009a) and in movies requires knowledge of 3,000 word families for 95% coverage and 6,000 and 7,000 word families for 98% coverage. Written genres such as novels and newspaper articles (Nation, 2006) are lexically even more demanding, requiring knowledge of up to 9,000 word families for 98% coverage. In addition, these results suggest that teachers select songs within the range of learners' vocabulary knowledge overall. Laufer (1998), for example, found that High School graduates in Israel have a vocabulary size of 3,500 word families, and Laufer (2001) showed that English majors in China have a vocabulary size of 4,000 word families. It seems, then, that teacher-selected songs can be considered lexically simple in comparison with other texts intended for a native speaker audience. In other words, songs can and do serve as an "entry level" authentic and unsimplified text genre that can be used as teaching material in the classroom and that overall meets the needs of language learners in terms of vocabulary demand.

However, this conclusion regarding the lexical demand of lyrics needs to be qualified: A separate analysis of three subcorpora categorised according to the proficiency level of the intended target learners revealed that there was little difference in terms of the overall vocabulary load between songs intended for use with beginners, intermediate and advanced learners. And an analysis of the 95% confidence interval for a sample of 108 songs (comprising an equal number of songs for beginners, intermediate learners

and advanced learners) indicated that 4,000 words is indeed a good estimate of the lexical demand of a majority of songs in the corpus. With required knowledge of 4,000 word families to reach 98% coverage, the vocabulary load in the song corpus might be at or even below the level of vocabulary knowledge of advanced learners. At the same time, this vocabulary load appears to be fairly high when considering the use of these songs with lower-level learners. Milton and Meara (1998), for example, showed that 15-yearold High School students in Greece and in Germany have a vocabulary size of 1,680 and 1,200 word families respectively.

How can this relatively high vocabulary load of songs used with lower-level students be explained? Firstly, it is possible that the results of the song corpus analysis are somewhat inflated due to the small number of words per individual song and the high number of repetitions of only a few words in one or very few texts. The 25 repetitions of the word *submarine* alone, for example, account for 14% of tokens in the 180-word song *Yellow Submarine* by The Beatles. The 25 repetitions of the word *yellow* account for another 14% of running words. More importantly, however, this comparatively high vocabulary load is in accord with the way teachers utilise songs in class. While the threshold of 98% coverage for adequate comprehension is intended for extensive reading and listening, for "reading for pleasure" (Hirsh & Nation, 1992) and for incidental vocabulary learning (Webb & Rodgers, 2009a), the teacher survey (Chapter 3) showed that inside the classroom songs are used for very specific purposes, including intentional, rather than incidental vocabulary learning. Over two-thirds of informants used songs purposefully to introduce new vocabulary, while 61.8% utilised lyrics to practise familiar vocabulary, and still over half of informants used songs to teach multi-word units.

In addition, songs used in the classroom are typically accompanied by a number of activities that aid comprehension and, thus, allow for a higher vocabulary load in the target text. According to the teacher survey, typical activities are gap-fill activities, discussions, and sing-alongs, all used by more than 70% of survey respondents. Moreover, the survey showed that informants typically repeated a song two to three times during a lesson. It seems, then, that teachers provide a great deal of support to aid learners' comprehension and facilitate vocabulary uptake. Overall, while the present methodological approach, that is, the analysis of vocabulary knowledge necessary to
reach particular coverage thresholds, allowed for a good estimation of the overall vocabulary load of lyrics, the findings should be interpreted with caution in regard to the unassisted comprehension of songs. It should always be kept in mind that, according to survey informants, teachers provide various forms of support for comprehension, which would permit a greater number of unknown words to occur in the target text. As Stæhr (2009) highlighted, necessary lexical coverage always depends on a number of factors including the purpose of an activity and the degree of comprehension required.

Apart from the lexical demand of songs, the present study was also concerned with the vocabulary learning opportunities that lyrics afford. As discussed above, the vocabulary load of teacher-selected songs appears to be fairly low when considering their use specifically with learners of advanced language proficiency. For example, knowing 4,000 words, as, for example, English majors in China do (Laufer, 2001), provides coverage of 98% of teacher-selected songs, which means that in a song of average length only five to six tokens are unknown. The analysis also showed that the song corpus contained little academic vocabulary. That is, the Academic Wordlist (Coxhead, 2000) accounted for only 0.33% of tokens in the corpus. This is not surprising given the fact that clearly fewer respondents reported using songs in classes with a focus on academic or specific professional purposes compared to courses for general or everyday language use. Additional research is required to investigate whether pop songs in general, rather than teacher-selected songs, display a similarly low rate of academic language.

Low-frequency vocabulary, that is, words from the BNC wordlists 9 to 20, account for only 0.56% of tokens in the corpus. It can, thus, be assumed that the present corpus also contains few words relevant for learners focussing on the vocabulary used in particular professional or academic fields. On the other hand, an analysis of the standardised typetoken ratio of the corpus indicates a low lexical diversity and a high degree of repetition. Teacher-selected songs are thus clearly less diverse and more repetitive than, for example, both the written and even the spoken part of the ICE-GB, the International Corpus of English (cf. Kreyer & Mukherjee, 2007). What is even more interesting is that the vocabulary in the pedagogical song corpus is also more repetitive than the language in the Giessen-Bonn Corpus of Popular Music (GBoP). We recall that the GBoP can be described as a corpus representative of pop music in general rather than of songs

selected on the basis of a pedagogical rationale (Kreyer & Mukherjee, 2007). This finding provides evidence that teachers consider the vocabulary load and the lexical diversity when selecting songs for in-class use.

Regarding the lexical learning opportunities afforded by teacher-selected songs, these findings indicate that these songs primarily provide opportunities to introduce low- and mid-frequency vocabulary. In addition, the predominant use of high-frequency vocabulary, the presentation of vocabulary in context and the high repetition of words in the lyrics suggest that songs, particularly when used with higher-level learners, can be useful to consolidate already familiar words and to facilitate the "deepening" of word knowledge, that is, the acquisition of word knowledge beyond a first form-meaning connection including grammatical functions, collocations and associations. Further research is required to establish whether this lexical profile is typical for pop songs in general or whether it is merely reflective of the teaching practices conveyed by respondents to the questionnaire in Study 1 (Chapter 3), where a majority of informants reported using songs to teach familiar vocabulary.

Finally, the present corpus study analysed additional characteristics of songs used for language-teaching purposes. The results confirm Murphey's (1990a) claims that songs are short and vague regarding personal referents and places referred to in the lyrics. The typical song found in the corpus contains 289 running words and has a length of around 3:50 minutes. This can be considered short, particularly when compared to other authentic text genres. Webb and Rodgers (2009b), for example, found that the shortest text in their corpus of TV programmes was the script of a news story containing a total of 523 words. Interestingly, the average lyrics in the pedagogical song corpus are also clearly shorter than the average song in the GBoP, which can be assumed to comprise 398 words as the GBoP contains 176,000 words derived from 442 songs. The most important difference between the two corpora is the "language-pedagogical slant" (Kreyer & Mukherjee, 2007, p. 31) of the current corpus, while the compilation of the GBoP was not guided by the pedagogical use of the target songs. This difference can be seen as an indication that language teachers take the length into consideration when selecting a song for pedagogical purposes.

Overall, the short length of pop songs compared to other authentic text genres can be considered an advantage for three reasons: Firstly, teachers do not have to spend an excessive amount of time assessing a song including its lexis. Secondly, songs can be used in their entirety during a language lesson due to the low number of words as well as the short time it takes to play them. And thirdly, the short length renders songs suitable for repeated listening and singing during one lesson, which can further support the entrenchment of word knowledge and phraseological patterns. Finally, as Murphey claimed, songs seem indeed to be vague: Proper nouns account for only 0.62% of tokens in the corpus and, thus, contain clearly fewer specific referents identifying particular people and places.

4.7 Limitations of the corpus study

The song corpus study has a number of limitations. Firstly, it should be noted that the present song corpus was small, as 635 songs only comprised 177,384 running words. Secondly, the low number of only 36 songs intended for use with advanced learners might have produced analysis results indicating smaller required vocabulary knowledge than could be expected if the three subcorpora had been more balanced in size. In addition, it seems that the frequent repetition of individual words in one or only few of the very short texts in the corpus might have resulted in a somewhat inflated estimation of vocabulary knowledge necessary to reach 95% and 98% coverage of the complete corpus as well as of individual songs. Further research is needed which takes into account the range and dispersion of words in the corpus. Thirdly, as elaborated in Chapter 2 (section 2.2.4), the methodological approach used in the song corpus analysis carries certain limitations in the present context. Above all, it has not been established whether coverage levels of 95% and 98% are indeed an appropriate threshold for the adequate comprehension of song lyrics. In addition, an analysis by means of the RANGE programme (Heatley et al., 2002) and Nation's 20 BNC wordlists (Nation, 2004, 2006) is intended to measure vocabulary knowledge necessary for unassisted reading and listening comprehension and incidental vocabulary learning. This purpose contrasts with the way questionnaire informants frequently reported using songs in class. As survey responses in Chapter 3 indicated, songs are frequently used for intentional teaching of new and familiar words and phrases, and comprehension as well as vocabulary

acquisition are typically assisted in a variety of ways, including the repeated playing of a song, the provision of written lyrics and complimentary activities, all of which affect the vocabulary knowledge learners need to bring to the task. However, as elaborated in Chapter 2, research results indicate that song lyrics display aspects of both written and spoken forms of discourse. It can, thus, be assumed that a coverage threshold for the comprehension of songs is located within a range of 95 to 98 percent of running words. Still, these methodological caveats should be kept in mind when drawing conclusions regarding the vocabulary knowledge required for the comprehension of songs.

In addition, the estimation of required vocabulary size is based on word families rather than individual words and builds on the assumption that learners have passive knowledge of all members of a word family, if one of them is known and if learners possess knowledge of inflectional and derivational affixes. However, it cannot be assumed that learners who know one member of a word family necessarily also know, or can deduce, other related words (Schmitt & Boyd Zimmerman, 2002; Schmitt, 2008). Using word families rather than individual words to gauge the vocabulary load of songs might, therefore, have resulted in a somewhat distorted estimation of the lexical demand of and the vocabulary learning opportunities afforded by songs.

At this point, it should also be reiterated that the focus of the present study was on songs preselected by teachers and material designers rather than on pop songs in general. Analysis findings can, thus, not be generalised to the genre of popular songs as a whole. The high number of songs in the corpus might allow for the cautious conjecture that pop songs overall tend to have a lower vocabulary load than other authentic and unsimplified text genres. However, this tentative claim needs to be confirmed by further research into the lexical profile of pop songs overall. Finally, it needs to be acknowledged that songs' vocabulary load is not the only aspect that teachers must and do consider when selecting songs for use in class. In fact, lexical demand is only one of many factors that teachers consider when choosing a song as teaching material. While the survey study in Chapter 3 highlighted some of these aspects, such as appropriate grammar, pronunciation and the popularity of a song with the target audience, the present corpus study focused exclusively on the vocabulary load of the lyrics.

4.8 Conclusion

As the present song corpus study indicates, popular English songs used in EFL and ESL classrooms around the world tend to be short, repetitive and lexically simple compared to other authentic, unsimplified text genres. It appears, then, that they can be seen as "entry level" authentic texts and as appropriate teaching material for the use with language learners in general. The predominant use of high-frequency words, the presentation of vocabulary in context and the high repetition of words in the lyrics suggest that songs, particularly when used with advanced learners, are suitable particularly for the consolidation of already familiar words and to support the acquisition of deep word knowledge beyond a first form-meaning connection, including grammatical functions, collocations and associations. This finding is in accordance with teaching practices reported by teachers in Study 1, which revealed that a majority of informants reported using songs to teach familiar vocabulary. However, prior to drawing any further conclusions in Chapter 6 regarding the recommended use of songs particularly for lexical learning, Chapter 5 presents a quasi-experimental intervention study, which assessed a possible beneficial effect of song-based learning on the retention of words and phrases. The description of a typical song used by English teachers provided by the present corpus study has informed the selection of a song for the intervention in the quasi-experimental study presented in Chapter 5.

Chapter 5 The quasi-experimental intervention study

5.0 Introduction

Research in the field of cognitive psychology has provided evidence that listening to a song, compared to listening to a spoken text, can result in superior retention of unconnected wordlists (Chazin & Neuschatz, 1990; Rainey & Larsen, 2002), and of longer connected text (Wallace, 1994; McElhinney & Annett, 1996; Calvert & Tart, 1993; Purnell-Webb & Speelman, 2008). However, a small number of studies have also produced contradictory results and concluded that a musical presentation of text is not beneficial (Kilgour et al., 2004) or can even be detrimental (Racette & Peretz, 2007) to verbal learning and verbatim retention. While some studies have included a beat-enhanced spoken format (Wallace, 1994; Purnell-Webb & Speelman, 2008), little research has been done to investigate the effect of a poetic text format compared to a prose text. Tillmann and Dowling (2007), however, found that reading and listening to a poetic text can result in increased recollection of structural surface features compared to a prose text. All of these studies have focused on native speakers and their memory for native language words and text. Only few studies have investigated the effect of song and poetry on vocabulary and text retention by second language learners. Overall they found that songs are at least not detrimental (Medina, 1990) and can even aid retention of short phrases (Ludke, 2010) and longer text (Hahn, 1972; Smith Salcedo, 2002). However, no previous study has looked at the use of songs in an ecologically valid classroom setting informed by actual teaching practices. Therefore, the current study was intended to lay some groundwork and investigate whether songs as they are currently used in many language classrooms around the world can benefit memory of verbatim text and, thus, aid second language acquisition. For this reason, the intervention used during the quasi-experiment described in this chapter was informed in detail by the two previous studies, the teacher questionnaire discussed in Chapter 3 and the corpus analysis discussed in Chapter 4. Teachers responding to the questionnaire frequently reported the use of songs to introduce new words, to practise familiar words and to teach multi-word units. In the pedagogical literature, songs are also often seen as suitable for vocabulary learning, including the acquisition of longer and idiomatic expressions, because both native and

non-native speakers seem to remember lyrics easily and over long periods of time (van Asselt, 1971; Abbott, 2002). Richards (1969), for example, claimed that "sometimes a structure or sentence pattern can be fixed in the mind of the learner through a song" (p. 162). Consequently, the focus of the present study was on retention of vocabulary in context and on the memorization of longer connected text. Verbatim memory or, in other words, the word-for-word recollection of text was chosen as a measure in this quasi-experiment. Verbatim memory is a form of episodic memory, that is, memory of previous learning events rather than memory of decontextualised knowledge, and was tested by means of a free recall test, a cued recall test and a recognition test (Baddeley, Eysenck, et al., 2009).

5.1 How can episodic memory aid vocabulary learning?

Episodic memory can benefit the memorization and retrieval of verbal information in a number of ways. First of all, it has been argued that episodic memory forms the basis of semantic memory, which includes decontextualised word knowledge (Tulving, 1972). Baddeley (1988) proposed that semantic memory is information abstracted from learning events initially processed and stored as autobiographical memory. Strong episodic memory can therefore be considered essential for the acquisition of decontextualised word knowledge. Secondly, the storage of information about an event as episodic memory allows for mental time travel (Tulving, 1985, 2002), that is, the mental re-visiting of the original learning event. Mental time-travel might assist the memorization and retrieval of words because the learner can re-experience the learning event including the verbal input encountered during that experience. As discussed in the literature review in Chapter 2, experimental participants used episodic memories such as the visualization of their own kitchen to produce decontextualised information, for example a list of objects typically found in a kitchen (Greenberg et al., 2009). Thirdly, episodic memory binds a variety of information from the learning event, such as emotions, movements and sensory experiences, which can serve as additional and interconnected memory traces during encoding. These different memory traces can also serve as effective memory cues during retrieval, triggering the recollection of the original learning event and its verbal content.

As discussed in Chapter 1, remembering strings of words is helpful because learners can recognise and also reproduce them as complete, unanalysed chunks. Producing holistic chunks of words rather than stringing together individual words is cognitively less demanding and gives speakers a more native-like fluency in language comprehension and production (Pawley & Syder, 1983). In addition, producing chunks "copied" from native speakers allows learners to produce more native-like output (Pawley & Syder, 1983). The memorization of phrases and text can also provide learners with models of word use in context. Rather than learning individual words in isolation, learners internalise a blueprint of word usage, including such important aspects of word knowledge as grammatical, collocational and pragmatic information.

5.2 Research question

The quasi-experimental intervention study set out to answer the following research question within an ecologically valid classroom setting: Does the presentation of a text in the format of a song result in higher verbatim text retention compared to a presentation in the format of a poem or prose text?

5.3 Methodology

The quasi-experiment overall comprised three learning conditions – a song condition, a poem condition and a prose condition. As research with native speakers indicated, a poetic spoken text can enhance retention of surface features of a text compared to a prose text (Tillmann & Dowling, 2007). Research in the field of second language acquisition has provided evidence that structural regularities of the text such as alliteration and assonance can benefit the retention of word strings (Boers & Lindstromberg, 2005; Lindstromberg et al., 2014; Boers et al., in press). In addition, the survey study discussed in Chapter 3 provided evidence that language teachers consider the rhythm in songs and also the rhyme in the lyrics to be relevant factors rendering songs suitable for vocabulary teaching. The value that survey-respondents attributed to rhythm and rhyme and the relatively lower importance ascribed to the melody highlight the importance of examining the effect of a poetic text presented without melody. All experimental groups received a similar language lesson that differed in two respects: (1) the format of the treatment text they worked with (a song, a poem or a prose text), and

(2) the oral production task the subjects participated in at the end of the lesson (a singalong or read-along). The difference between learning conditions was thus based on the different forms of presentation of the target text. The lesson was immediately followed by a battery of three memory tests measuring different aspects of verbatim text retention (two recall tests and one recognition test). These tests were again repeated one week later. Figure 5.1 provides a broad overview of the intervention and subsequent test battery. The details of the intervention and the testing will be further discussed below.

	SONG GROUP	POEM GROUP	PROSE GROUP				
TEXT FORMAT	song	poem	prose				
INTERVENTION	typical language lesson						
ORAL PRODUCTION	sing-along read-along read-along						
POSTTESTS (immediate, delayed)	free recall, cued recall, recognition						

Figure 5.1: Overview of the quasi-experimental intervention and post-test battery

5.3.1 Participants

The quasi-experimental intervention study was conducted at three different institutions in three European countries: Serbia, Germany and Belgium. The three institutions differed in various respects, most importantly regarding the number and the proficiency levels of participants and the number of experimental groups. The institution with the largest number of participants was a university in Serbia. 63 undergraduate students (49 female, 14 male) of translation studies (Serbian-English) participated in the study. The selection of experimental groups was based on existing classes, which resulted in an unequal distribution across groups: There were 13 participants in the song group (4 male, 9 female), 18 in the poem group (2 male, 16 female) and 33 in the prose group (8 male, 24 female). The participants' average age was 21.1 years. The professors at the university assessed this cohort as having an advanced level of proficiency in English. Due to miscommunication between the researcher and the administration at the institution, the song group was not available for delayed post-testing. Consequently, only delayed data from the poem group and the prose group were available for analysis. The second participating institution was located in Germany: A total of 18 secondary students (9 male, 9 female) in two existing classes participated in the study. As the selection of experimental groups was based on existing classes, only two of the three learning conditions (song, prose) could be implemented. The song group comprised 8 participants (4 male, 4 female), while the prose group comprised 10 subjects (5 male, 5 female). The participants' average age was 16.6 years, and their English teacher assessed their overall proficiency level as low-intermediate. The third participating institution was a secondary school in Belgium. A total of 24 participants (14 female, 10 male) in two existing classes participated in the study. Again, the number of experimental groups was determined by the number of existing classes. Consequently, only two learning conditions (song, poem) could be compared at this institution. The song group comprised 9 participants (5 female, 4 male), and the poem group comprised 15 participants (9 female, 6 male). The participants' average age was 17.5 years. The participants' English teacher assessed their overall proficiency level to be high-intermediate.

5.3.2 Materials used during the intervention

The intervention involved listening to and reading a text in one of three formats. The intervention was preceded by a pre-teaching activity to familiarise participants with 20 expressions from the text. During the treatment lesson, participants engaged in a gap-fill activity as well as two small-group activities. All of these tasks required printed or audio-recorded material, which I will discuss in detail below.

5.3.2.1 The target texts in their written and audio-recorded versions

During the intervention, the participants both listened to and read a target text three times as a song, a poem or a prose text. The three text formats were based on the lyrics of a song taken from the pedagogical corpus discussed in Chapter 4. The selected song was Nickelback's "If today was your last day" (Kroeger, 2008), in which a lyrical subject muses about living life to the fullest and each day as if it was the last. This topic lent itself to a meaning-focused discussion without requiring any background knowledge and further elaboration of the topic. More importantly, the song met a majority of the criteria of a typical song used in the English classroom as described in the corpus study. The song also contained a high number of formulaic sequences, that is, collocations, idiomatic expressions and phrasal verbs. As questionnaire respondents referred to the teaching of multi-word units as a frequent teaching goal and as teaching of fixed expressions was often the stated purpose of song use in pedagogical articles (Abbott, 2002; Stansell, 2005; Engh, 2013; Kerekes, 2015), it was seen as essential that the target song contained a certain number of formulaic expressions.

"If today was your last day" was published comparatively recently (2008) and was produced and performed by a North-American band with a male lead singer. The abridged version (3:20 minutes) used during the experiment was within the typical length of songs used by language teachers. It was based on the slightly longer album version (4:06 minutes) and shortened for the purpose of using it during the intervention. In the abridged version, the song's long fade-out was eliminated because the continuous repetition of particular lines would have rendered both the prose version and the poem version unnatural. It was reasoned that teachers also frequently refrain from playing the fade-out of a song due to time constraints or because their students might lose interest during long fade-outs. In that case, the abridgement would be acceptable. The required knowledge of 3,000 word families for 98% coverage of the text (see chapters 2 and 4) places the song in the typical range displayed by songs in the pedagogical corpus in terms of lexical demand. The song might have contained words unknown to participants of intermediate proficiency. This was not considered to be a problem as any unknown words were clarified during the intervention. In written form, all experimental groups saw the texts only in their cloze version. See section 5.3.4 for more details on the gap-fill activity.

The song format

The song consisted of 292 words in two verses, a main chorus (repeated twice), an extension of the chorus, a lyrical bridge and a final additional repetition of the first line of the chorus. (See the written lyrics in form of a gap-fill activity in Appendix C.) The audio version of the song started with a six-second musical intro and contained an additional instrumental bridge spanning ten seconds between the first verse and the subsequent chorus. In written form the song was presented in a poem structure. That is, the chorus

and the verses were presented as separate stanzas which consisted of short poetic lines often ending on a rhymed syllable. The (abridged) audio version of the song was played to the participants using a CD and CD-player. The song was introduced to the participants as a song motivating listeners to live life to the fullest.

The poem format

The poem was identical to the song format in wording and written structure. However, the audio version differed in several respects: The text was read by a volunteer male native Canadian speaker whose voice was similar to the singer of the song version. The reading was recorded in a sound-reduced room at Victoria University of Wellington. The audio recording had a running time of 3:12 minutes and did not have musical accompaniment. The reading adhered to a similar pattern of pauses between stanzas and at the end of lines as the song. In lieu of a musical intro, the reading started with six seconds of noise. The noise consisted of the hustle and bustle of an audience slowly calming down before the beginning of a performance, however without any actual discernible words. This noise provided clues about the beginning of the reading, equivalent to the musical introduction of the song format. In lieu of the musical bridge found in the song, the poetical reading contained a break of the same length, filled with the sound of turning pages, which allowed for possible cognitive processes such as rehearsing, encoding or forgetting to occur, just as they might in the song condition. However, there needed to be noise that gave listeners a clue that they had not reached the end of the recording and that further text was to come, as the song did. The poem was introduced to the participants as an original poem motivating readers to live life to the fullest. Under the poem condition, no mention was made that it was the text of an original song.

The prose format

The layout of the prose text was a running text, which contrasts with the poem structure of the song and the poem format (see a number of sample lines in Figure 5.2). However, the stanzas were also presented as separate paragraphs. The following modifications were made to the prose text:

• the addition of six words (Then he asked; so; and; and);

- the deletion of one word (*and*);
- the changed order of clauses in one sentence to eliminate a clear rhythmic and rhyme pattern (*You know, regardless of who you are, it is never too late to shoot for the stars*).

The modifications were necessary in order to produce a convincing and natural prose text, and to disrupt a clearly perceivable rhythmic pattern. The changes are highlighted in Figure 5.2.

Figure 5.2: An example of the prose text with highlighted differences to the other formats

Against the should be a way of life. What's worth the prize is always worth the						
fight. Every second 'cause there is no second try. So live it like you'll never live						
it twice. Don't take the free ride in your own life.						
So: If today was your last day and tomorrow was too						
goodbye to yesterday? Would you each moment like your last, leave old						
pictures in the past, every dime you have? Would you call old friends you never						
see <u>and</u> on memories? Would you forgive your enemies? Would you find that						
one you're dreaming of <u>and</u> swear up and down to God that you'd finally fall						
love. If today was your last day?						

These changes resulted in a total number of 297 tokens, five more than in the two other formats. The audio version of the prose was recorded in a sound-reduced room at Victoria University and had a running time of 3:11 minutes. It was read by the same male Canadian volunteer as the poem. The prose text was read at a similar rate as the song and included pauses of similar length in the same places as the other two formats. The same recording patterns were followed for the prose text as for the poem (see above). The prose text was introduced to the participants as a funeral speech in a movie that was intended to encourage the listeners to live life to the fullest. No mention was made of the song or poem.

5.3.3 The intervention – a typical language lesson

The intervention consisted of a typical language lesson of slightly over 30 minutes. The only differences in the design of the teaching units for the different experimental groups

were the text format and the oral-production activity. The treatment lesson was informed by language instructors' cognitions and teaching practices elicited by means of the international teacher questionnaire (discussed in Chapter 3). The information provided by the questionnaire informants shaped all aspects of the treatment lesson including its length, the number of repetitions of the target text, the choice of material and the activities that participants engaged in. Figure 5.3 has a short description and timing of each element of the intervention and the material used.

TITLE	DESCRIPTION	MATERIAL	TIMING (MIN)
Information and consent	Participants silently read through the information sheet and are encouraged to ask questions. If they consent to the use of their data as detailed in the consent form, they sign and return the form to the researcher.	information sheetconsent form	5
Introduction	The researcher provides a general introduction to the lesson with a short preview of what activities (listening, gap fill, group discussion etc.) the students will participate in. The purpose of the study is not mentioned.		1
Pre-teaching of target vocabulary	The participants receive a list of target words and expressions and a translation in their native language. They silently read through the list and familiarise themselves with the expressions to facilitate listening and reading comprehension of the target text.	 list of English target items 	2
Listening to/reading of target text (1) and gap- fill exercise	Participants listen to a recording of the song, poem or prose text. While listening they fill out a rational cloze-version of the text.	 audio-recording of song, poem or prose text rational cloze- version of the text 	5
Researcher-led discussion of correct answers for the gap-fill exercise	The researcher/teacher asks participants for the correct answers to fill the blanks. The participants call out the correct answers upon request and the researcher writes them on the white board.	• (filled out) cloze- version of the text	3

Figure 5.3: Description of the treatment procedures (including material and timing)

Listening to/reading of text (2) to verify correct answers in the gap-fill exercise	Participants listen to the recorded text a second time. They read along, verify their answers and check that they are now able to hear the correct answers.	 audio-recording of song, poem or prose text (filled out) cloze- version of the text 	4
Introduction to small group activities and teacher-fronted follow- up	The researcher/teacher explains the activities that are about to follow. The participants form small groups of three.		2
Small group activity: comprehension question	Participants receive a worksheet with instructions for two tasks and space to write answers. The first task involves finding and writing down five statements from the target text. After 2.5 minutes, the researcher encourages the participants to begin the second task.	 (filled out) cloze- version of the text worksheet with instructions for two tasks 	2.5
Small group activity: personalised topic discussion	The second task requires the participants to come up with a top 5 list of things they would do if today was their last day. The participants need to discuss the matter and agree on a list that that they can all agree on.	 (filled out) cloze- version of the text worksheet with instructions for two tasks 	2.5
Teacher-led class presentation/discussion of results from the small-group activities	The teacher/researcher prompts the participants to present their answers to the entire class. The teacher/researcher and the participants engage in short discussions about the answers.	 (filled out) cloze- version of the text worksheet with instructions for two tasks 	2 (task 1) 3 (task 2)
Listening to/reading of text (3) including a choral oral production of the text (sing-along or choral reading)	<u>Sing-along</u> : The participants and the researcher/teacher sing along with the audio-recording of the song (3), reading from their written copy of the lyrics. - OR - <u>Read-along</u> : The participants and the researcher/teacher read along (aloud) with the audio-recording of the poem/prose text.	 audio-recording of song, poem or prose text (filled out) cloze- version of the text 	4
Collection of all materials	The researcher/teacher collects all materials and wipes the answers for the cloze activity off the whiteboard.		0.5

5.3.3.1 Information and consent

This study was approved by the Human Ethics Committee at Victoria University of Wellington on August 15, 2012. All documentation can be found in Appendix C. One student in Germany did not consent to the use of his test results, so his data was destroyed immediately following data collection.

5.3.3.2 The experimental activities

The four most popular activities used with songs as reported by questionnaire informants (Chapter 3) were integrated into the treatment lesson. The activities were a gap-fill task, small group work which required responding to a comprehension question, a personalised topic discussion, and a sing-along (or read-along). In addition, all students engaged in pre-teaching of target expressions before participating in these activities.

The pre-teaching activity

18 multi-word units including phrasal verbs, idioms and collocations were brought to the participants' attention by means of a pre-teaching activity, during which the participants received a list of expressions in English and a translation in their native language. These items were selected based on the following characteristics: They were fixed multi-word units consisting of two or more highly frequent words which, based on their frequency, were likely to be known by the participants. 87% or 54 of the individual words found in these multi-word units can be found in the first 1,000 words of the Nation's (2004, 2006) frequency lists based on the British National Corpus (BNC), 10% or a total of 6 words can be found in the second BNC-frequency band, and only 3% or exactly 2 words are found in the third 1,000 words of Nation's frequency bands. Figure 5.4 shows the list of expressions and translations that the participants worked with. Note that the actual worksheets only contained the native language of the particular cohort. The Dutch translation was provided by the researcher's supervisor, while the Serbian translation was produced by a professor of translation studies (Serbian – English) at the participating Serbian university. The German translation was produced by the researcher herself and checked by a native speaker of German/English language instructor at a university in Germany. The participants silently read through the list and familiarised themselves with the expressions for two minutes.

Figure 5.4: Expressions from the target text and their translations in the participants' native languages presented in a pre-teaching activity

English expressions	German	Dutch	Serbian
1) to give someone advice	jemandem einen Rat geben	iemand raad geven	Posavetovati nekoga
2) a given right	ein Anrecht	een gegeven recht	Pravo
3) to leave no stone unturned	nichts unversucht lassen	geen middel onbeproefd laten	Prevrnuti svaki kamen
4) to take the path less travelled by	wörtlich: den kaum genutzten Weg nehmen; übertragen: Neues, Ungewöhnliches wagen	zijn eigen weg gaan	Izabrati teži put, ići trnovitim putem
5) to take the first step	den ersten Schritt machen	de eerste stap zetten	Napraviti prvi korak
6) against the grain	gegen den Strom; gegen den Strich	tegendraads	(Plivati) protiv matice
7) Every second counts.	Jede Sekunde zählt.	Elke seconde telt.	Važan je svaki minut.
8) to take the free ride	wörtlich: ein Trittbrettfahrer sein; übertragen: sich das Leben leicht machen	het zich gemakkelijk maken	Dobiti nešto na gotovo
9) to dream of something	von etwas träumen	ergens van dromen	Sanjati o nečemu
10) God above	Gott im Himmel	God in de hemel	(Sam) Bog
11) to fall in love	sich verlieben	verliefd worden	Zaljubiti se
12) to make one's mark by	sich einen Namen machen durch; sich profilieren durch	bekend of beroemd worden voor iets	Istaći se
13) to mend a broken heart	ein gebrochenes Herz heilen	iemands liefdesverdriet genezen	Izlečiti slomljeno srce, zaceliti rane
14) regardless of	ungeachtet; ohne Rücksicht auf	ongeacht, niettegenstaande	Bez obzira na
15) to do whatever it takes	alles Menschenmögliche unternehmen; alles Erdenkliche tun	koste wat kost	Upotrebiti sva sredstva
16) Let nothing stand in your way.	Lass dich durch nichts aufhalten.	Niets in je weg laten staan.	Ukloniti sve prepreke.
17) the hands of time	die Zeiger der Zeit (poetisch)	bij verloop van tijd	Kazaljke na satu, vreme, zub vremena
18) to be on someone's side	auf jemandes Seite sein; jemanden unterstützen	aan iemands kant staan, iemand steunen	Biti na nečijoj strani

The gap-fill activity

The teacher questionnaire indicated that the gap fill is the activity most commonly used in a song-based lesson. The rational cloze-text contained 22 gaps targeting 17 different words. Ten of these words formed part of the expressions previously highlighted during the pre-teaching activity (*given*, *unturned*, *grain*, *counts*, *above*, *in*, *make*, *shoot*, *of*, *hands*), while six words were found outside these highlighted expressions (*late* [3X], *say* [3X], *live* [2X], *donate* [2X], *reminisce*, *rewind*). The rational cloze-text can be found in Appendix C.

The comprehension question

The participants engaged in small-group work to respond to the following comprehension question: "The singer/speaker tries to encourage us to live life to the fullest. What are 3 things he suggests we could do?"

The topical discussion

The participants engaged in a short discussion in small groups relating the topic of the target text to their personal life. The discussion was guided by the following question: "What would you do, if today was your last day? Discuss in your group and agree on a list of 5 things you would do. Take notes so you can present the results to the class."

The sing-along and the choral reading activity

The survey showed that over 60% of informants included one or two sing-alongs in their song-based lessons. One sing-along was, thus, integrated into the experimental lesson. The equivalent for the non-melodic texts consisted of a read-along, a form of choral reading. Choral reading and echo reading are still part of foreign language teaching (McCauley & McCauley, 1992; Peregoy, Boyle, & Cadiero-Kaplan, 2008). For the sing- or read-along, participants listened to the audio-text for a third time. They were told to read along on their worksheets and either sing or speak out loud. The read-along was implemented to engage the poem group and the prose group in an oral, productive task equivalent to the singing. The participants in the poem groups and prose groups were encouraged to imitate the pacing, prosody and dramatic performance of the speaker.

5.3.4 The test instruments

Verbatim memory or, in other words, the word-for-word recollection of text was chosen as the dependent variable in this quasi-experiment. Verbatim memory of the target text was measured immediately following the intervention by means of three direct memory tests, "in which the instructions … make reference to the target event" (Richardson-Klavehn & Bjork, 1988, p. 477). Direct memory tests require the retrieval of elements of that past event and thus measure episodic memory. In the present case, they required verbatim memory. That is, they measured the retention of the exact wording of the treatment text as experienced during the learning session. Verbatim memory is a form of episodic memory, that is, memory of previous learning events rather than memory of

decontextualised knowledge, and was tested by means of a free recall test, a cued recall test and a recognition test (Baddeley, Eysenck, et al., 2009). While participants were warned that the lesson would be followed by a number of tests, they were unaware that these tests would measure verbatim text retention. The participants first took a freerecall test, followed by a cued-recall test and finally a recognition test. These three tests measured retention of verbatim text by gauging the participants' ability to produce the orthographic form of words and phrases of the target text (free-recall test, cued-recall test) or by measuring recognition memory of the written surface form of phrases from the target text (recognition test). The three memory tests differed in cognitive demand, as productive recall is considered to be more difficult than receptive recognition (Baddeley, Eysenck & Anderson, 2009). In addition, the quantity of the provided retrieval cues affected the tests' difficulty. While the free-recall test provided no cues apart from the first phrase of the target text, the cued-recall test provided 50% of the words as potential retrieval cues. The recognition test provided 18 entire original phrases, each alongside three incorrect distractors. Both the original phrase and the distractors served as retrieval cues with the potential to trigger a memory of the learning event and the learnt material.

The three tests did not only differ in difficulty or cognitive demand. They also differed in terms of their sensitivity to partial learning. It is understood that less information is needed to recognise an item (or reject a distractor) than to recall an item (Postman, 1963; Laufer & Goldstein, 2004). The recognition test was, thus, included to capture partial memory of verbatim text and to show possible differences in this regard between the three experimental groups. Furthermore, using three tests that differed in the quantity as well as quality of retrieval cues provided addressed a possible accessibility problem. As Tulving and Pearlstone (1966) explained, a failure to recall information does not necessarily imply forgetting, in other words, an availability problem, but possibly "represents a failure to 'find' otherwise intact traces in storage" (p. 381). Retrieval cues aid participants in accessing the retained information, thus solving a possible accessibility problem. The cued-recall test is, thus, more likely than free recall to measure actual retention of information rather than the ability to retrieve – potentially retained – information. A recognition test provides even stronger retrieval cues, as it comprises

entire phrases as well as a number of distractors of similar semantic content. Cued recall and recognition, thus, differ in regard to the strength of the provided retrieval cues. As explained above, they also measure different types of information: While cued recall measures the retention of information "sufficient for the reproduction of words" (Tulving & Pearlstone, 1966, p. 389), the recognition test measures recognition memory, which requires only partial learning.

During delayed post-testing, the participants repeated the three direct memory tests. While the participants knew that they would meet the researcher again one week later, they did not know that they would be tested again. The delayed posttests can thus be considered unannounced. While the immediate posttesting constitutes additional practice of the verbatim text and the target items, participants of all three groups (song, poem, prose) engaged in the same procedures and all had the same opportunity for additional rehearsal. At the end of delayed testing, participants responded to a questionnaire inquiring about their personal information, their views of the treatment lesson and the use of songs in the language classroom. These data were not included in this analysis.

5.3.4.1 The free-recall test

The free-recall test required participants to write down as many words as they remembered of the previously encountered treatment text. The participants had six minutes to complete the test. They were encouraged to write the words in the original order if possible. However, they were explicitly told that any order was acceptable. No memory cue apart from the first sentence was provided. The worksheet including full instructions can be found in Appendix C.

5.3.4.2 The cued-recall test

The cued-recall test followed the free-recall test. Participants had six minutes to complete the test. The test had the appearance of a gap-fill activity with 50% of words (111 in all three conditions) in each sentence replaced by lines, while 50% of words remained as possible retrieval cues. The text was not presented in its original layout, that is, in a poem structure (song and poem) or in a running text (prose text). Instead, each sentence was numbered and presented separately. The order of the original text was

retained in order to allow for sequential cueing and recall. The selection of words as retrieval cues was done randomly with the qualification that none of the words provided could be relevant in selecting the correct answer in the subsequent recognition test. The missing words were represented by (under)lines with one line representing one word, thus indicating the exact number of missing words. Figure 5.5 provides an example of the cued-recall test. The full worksheet including instructions can be found in Appendix C.

Figure 5.5: Excerpt from the cued-recall test



5.3.4.3 The recognition test

The third measure was a forced-choice recognition test. The participants were presented with 18 test items, each comprising a correct phrase from the treatment text alongside three distractors, that is, phrases that differed from the original in one or several respects. The participants were asked to identify the original phrase encountered previously during the learning session. Figure 5.6 provides one test item as an example.

Figure 5.6: Sample item from the recognition test

- A You know, it's never too late to reach for the stars regardless of who you are.
- B You know, it's never too late to shoot for the stars regardless of who you are.
- C You know, it's never too late to <u>aim for</u> the stars regardless of who you are.
- D You know, it's never too late to strive for the stars regardless of who you are.

The forced-choice design reduced the chance of correct guessing, which is a higher risk in yes-no recognition tests. The location of the original phrase among three distractors for each test item was determined randomly with the aid of a random number generator (www.random.org). The participants had five minutes to complete the test. The

worksheet including instructions can be found in Appendix C. The distractors contained one to two changes from the original text. The differences were underlined in order to be easily detected by the participants. As the focus was on verbatim-text retention, that is, retention of surface-form rather than meaning or gist, care was taken that the semantic content of the distractors did not differ greatly from the original target phrase. While the changes in the wording of the incorrect alternatives affected both open-class and closedclass words, the number of changes affecting content words was greatly reduced following the pilot study (discussed below).

5.4 Data analysis and test scoring

Following McElhinney and Annett (1996), the results of the free-recall test were scored by counting each correctly recalled word. Only correctly spelled words were counted. Word order and context did not affect the score. Falsely remembered words were ignored rather than "punished", for example by subtracting points. Indications of partial memory such as for example the word *diamond* instead of the original *dime*, were ignored as they could not be objectively assessed and quantified. Total scores were calculated as a ratio of the total possible number of words because the prose text contained five tokens more than the other two text formats. The cued-recall test was scored by counting each word correctly recalled. Only correctly spelled words were counted, and only words written in the appropriate gap were accepted. However, within the gap any word order was considered acceptable. Total scores were again calculated as a ratio of the total possible number of words. The recognition test was scored by awarding one point if an original phrase was correctly identified, while no point was awarded if a distractor was selected. The statistical analysis of data was performed using SPSS, Version 19.0 (2010) and is described in greater detail below in the results section.

5.5 Pilot study

A pilot study was conducted at Victoria University of Wellington. Participants were Malaysian students in a 3rd-year applied linguistics class. All participants were scholarship holders pursuing a degree in TESOL on a 2-year exchange to New Zealand. In total, 40 students participated in the study, 12 in the song group, 16 in the poem group, and 12 in the prose group. However, participation in delayed posttesting dropped to 26. The pilot

study served to determine the appropriate timing for the treatment activities and the post-intervention testing and as a test-run of the three memory tests.

5.5.1 Implications of the pilot study for the main study

The pilot study revealed no significant differences between the three learning conditions (song, poem, prose) on all three memory tests. However, the song group and the poem group produced clearly higher mean scores on the delayed recall tests than the prose group. A one-way ANOVA comparing the mean test scores of the three learning conditions revealed a slight trend towards statistical significance during delayed free recall (p = .06) and delayed cued recall (p = .08). The piloting of the memory tests also highlighted a number of flaws in the test design. Firstly, the pilot study showed that the low number of cues in the cued-recall test resulted in an overall test performance that barely differed from free recall. As a consequence, the number of words provided as cues in the cued-recall test during the main study was increased to 50% of words per sentence, in order to render the test less demanding and to differentiate more clearly between availability and accessibility of the memorised text. Furthermore, it was found that several items of the recognition test were too easy and produced a ceiling effect. On the immediate recognition test seven test items (out of 18) were answered correctly by 75% or more of participants, with one phrase being correctly identified by 92.5% of participants. During delayed posttesting two items were answered correctly by over 75% of participants. Consequently, these test items were unable to capture any possible differences between learning conditions. Following Gurevich, Johnson and Goldberg's (2010) claim that content words are more easily remembered than function words, the number of test items targeting content words was reduced and the number of test items targeting function words was increased.

Other changes were related to the tests' formatting. In the pilot version of the cuedrecall test missing words were represented by one continuous line with the number of missing words indicated in brackets: _________(3). The pilot participants pointed out that this layout confused and distracted them. For the main study, the numbers were removed and each missing word was indicated by a separate line. That is: ________. The recognition test also underwent some changes to its layout. In its piloted version, the differences between the original phrase

and the distractors were not underlined, which effectively turned the recognition test into a search-and-find activity. In other words, the pilot participants spent a great deal of time and effort on searching and finding the differences between the sentences. As a consequence, in the main study the differences were underlined to simplify comparison and recognition.

5.6 Results

This section reports descriptive statistics for overall performance on the three direct memory tests (free recall, cued recall, recognition) at each of the three participating institutions, as well as inferential statistics comparing differences in performance across the three – or two – learning conditions (song, poem, prose) at each separate institution. To discover possible differences between learning conditions, I conducted a mixed within-between subjects repeated-measures ANOVA, which allowed for a collective analysis of the results of all three memory tests together. In addition to an overall analysis across three memory tests, further statistical analyses were conducted investigating an effect of the learning condition on the test scores of each memory test separately. This was done to determine whether possible differences between learning conditions varied depending on the type of test or whether all forms of memory, that is, free recall, cued recall and recognition, were affected in a similar way. Due to the fact that in all but one case only two experimental groups were available, inferential statistics for each test type separately were produced conducting pairwise comparisons (song/poem, song/prose, poem/prose). This was done by means of a T-Test or, if scores were not distributed normally, its non-parametric alternative, a Mann-Whitney U Test.

5.6.1 Results for each participating institution (Serbia, Germany, Belgium)

This quasi-experiment investigated the effect of three learning conditions on the participants' verbatim recall and recognition of a target text. The research question it was designed to answer was: Can the presentation of a text in the format of a song result in higher verbatim text recall compared to a presentation in the format of a poem or prose text? Figure 5.7 gives an overview of the available learning conditions in immediate and delayed posttesting at the three participating institutions.

Institution	Learning condition - Immediate -	Learning condition - Delayed -		
	song			
Serbia	poem	poem		
	prose	prose		
Germany	song	song		
Germany	prose	prose		
Belgium	song	song		
Deighann	poem	poem		

Figure 5.7: Overview of the	learning conditions	implemented at thr	ee institutions
-----------------------------	---------------------	--------------------	-----------------

Subjects at all three institutions participated in immediate and delayed posttesting. However, as explained in section 5.3.1, in Serbia the song group did not participate in delayed posttesting. In addition, only two classes were available in both Germany and Belgium, limiting the number of learning conditions to two at each of these institutions. As a consequence, in Germany, I only compared the learning conditions song and prose. In Belgium, I compared song and poem.

To discover possible differences between learning conditions, I conducted a mixed within-between subjects repeated-measures ANOVA, which allowed for a collective analysis of results of all three memory test together. For this purpose, the three memory tests were treated as a repeated measure. The variable "test type" was expected to yield significantly different results, as the three memory tests decreased in cognitive demand with free recall being the most and recognition the least demanding measure (see section 5.3.4). More importantly, the three tests applied a different scoring system in terms of total possible scores: Free recall had a total possible score of 220 for the song condition and the poem condition and 225 for the prose condition. Cued recall required verbatim recall of only 50% of the words of the original texts and, thus, had a total possible score of 111 in all three learning conditions. The recognition test had a total possible score of 18 points for 18 original phrases correctly identified. In this section, results are displayed as ratios in order to allow for a collective analysis by means of one statistical test only. The main interest was a possible overall effect of learning condition.

5.6.1.1 Serbia

In immediate posttesting in Serbia, three learning conditions – song, poem and prose – were compared. Table 5.1 shows the descriptive statistics for their performance on the

three memory tests in immediate posttesting across the three learning conditions. Results are provided as a ratio of the total possible scores, and raw mean scores are displayed in brackets.

Table 5.1: Serbia – performance on three memory tests during immediate posttesting for three learning conditions (number of participants [N], mean scores [M], standard deviation [SD])

	(4.5	Free Recal))	Cued Recall			F	Recognition		
	(to	tal: 220/2	25)		total: 111			(total: 18)		
	Ν	М	SD	Ν	М	SD	Ν	М	SD	
SONG	13	.3224	.0919	13	.5343	.1797	13	.7650	.0939	
5010		(70.93)			(59.31)			(13.77)		
DOEM	18	.2745	.1037	18	.4710	.1353	18	.6944	.1115	
FOLIVI		(60.39)			(52.28)			(12.50)		
DROSE	29	.2697	.0966	29	.4272	.1354	29	.6418	.0886	
FROJE		(60.68)			(48.70)			(11.55)		

In Serbia, the song group consistently produced the highest mean scores on all three tests, as can be seen in Table 5.1 above. The prose group consistently produced the lowest mean scores. A mixed within-between subjects repeated-measures ANOVA revealed the following: As expected, the variable "test type" was significant, with Wilks' Lambda = .057, F (2.000, 56.000) = 465.473, p = .000, partial eta squared = .943. There was no significant interaction between test type and learning condition, Wilks' Lambda = .881, F (4.000, 112.000) = 1.824, p = .129, partial eta squared = .061. A one-way ANOVA exploring differences between test scores averaging across the three test types showed a significant overall effect of learning condition, F (2, 57) = 4.158, p = .021, $\eta p^2 = .127$. A post-hoc pairwise comparison showed no significant difference between the learning conditions song and poem (p = .286) and poem and prose (p = .770). Only the learning conditions song and prose yielded significantly different results (p = .017).

In delayed posttesting in Serbia, only data from the poem group and the prose group were available. Table 5.2 shows the descriptive statistics for their performance on the three memory tests. Results are provided as a ratio of the total possible scores, and raw mean scores are displayed in brackets.

Table 5.2: Serbia – performance on three memory tests during delayed testing for two learning conditions

	Free Recall			Cued Recall			Recognition			
	(total: 220/225)			(total: 111)			(total: 18)			
	N	М	SD	N	М	SD	Ν	М	SD	
POFM	16	.2895	.1297	16	.5434	.1462	16	.7569	.1297	
		(63.69)			(60.32)			(13.62)		
PROSE	28	.2489	.1251	28	.4624	.1539	28	.6468	.1258	
		(56.00)			(52.71)			(11.64)		

As Table 5.2 shows, in delayed posttesting the poem group consistently produced higher mean scores than the prose group on all three tests. A mixed within-between subjects repeated-measures ANOVA revealed that, as expected, the factor "test-type" was significant, with Wilks' Lambda = .085, F (2.000, 41.000) = 222.015, p = .000, ηp^2 = .915. There was no significant interaction between test type and learning condition, Wilks' Lambda = .932, F = (2.000, 41.000) = 1.488, p = .238, ηp^2 = .068. A one-way ANOVA for test scores averaging across all three test-types revealed a significant overall effect of learning condition (poem, prose), F (1, 42) = 4.412, p = .041, ηp^2 = .095.

5.6.1.2 Germany

In immediate posttesting in Germany, only a song group and a prose group were available and the number of participants was low. Table 5.3 shows the descriptive statistics for their performance on the three memory tests. Results are provided as a ratio of the total possible scores, and raw mean scores are displayed in brackets.

Free Reca (total: 220/2			Free Recall (total: 220/225))	F	ecognitio (total: 18)	n
	Ν	М	SD	Ν	М	SD	Ν	М	SD
SONG	8	.1676	.0708	8	.2208	.1009	8	.5903	.1453
50110		(36.87)			(24.51)			(10.63)	
PROSE	10	.1178	.0578	9	.2446	.1051	9	.4753	.2099
INCOL		(26.51)			(27.88)			(8.56)	

Table 5.3: Germany – performance on three memory tests during immediate testing for two learning conditions (song, prose)

Table 5.3 shows that the song group produced higher mean scores during free recall and recognition. On the free-recall test, the song group on average produced 10.36 words more than the prose group. And on the recognition test, the song group on average

correctly identified over ten target phrases, while the prose group only correctly selected an average of 8.56 phrases. However, the prose group produced a higher mean score for cued recall. A mixed within-between subjects repeated-measures ANOVA revealed the following: As expected, the factor "test type" was significant with Wilks' Lambda = .095, F (2.000, 14.000) = 66.335, p = .000, ηp^2 = .905. There was also a significant interaction between test type and learning condition, Wilks' Lambda = .348, F (2.000, 14.000) = 13.122, p = .001, ηp^2 = .652. A one-way ANOVA exploring differences between test-scores averaging across the three test types showed no significant overall effect of learning condition, F (1, 15) = .749, p = .400, ηp^2 = .048.

In delayed posttesting in Germany, again only data from the learning conditions song and prose were available. Table 5.4 shows the descriptive statistics for their performance on the three memory tests. Results are provided as a ratio of the total possible scores. In addition, raw mean scores are displayed in brackets.

	Free Recall (total: 220/225)				Cued Recal)	F	Recognition (total: 18)	n
	Ν	М	SD	Ν	М	SD	Ν	М	SD
SONG	8	.1557	.1055	8	.2444	.1395	8	.5625	.1498
50110		(34.25)			(27.13)			(10.13)	
PROSE	9	.1052	.0514	9	.1832	.1226	9	.3951	.1533
TROOL		(23.67)			(20.88)			(7.11)	

Table 5.4: Germany – performance on three memory tests during delayed testing for two learning conditions (song, prose)

In delayed posttesting in Germany, the song group consistently produced higher mean scores than the prose group on all three direct memory tests. The difference seemed to be most pronounced on the recognition test, where the song group on average correctly identified three sentences more than the prose group. A mixed within-between subjects repeated-measures ANOVA revealed the following: As expected, the factor "test type" was significant, Wilks' Lambda = .075, F (2.000, 14.000) = 86.324, p = .000, ηp^2 = .925. There was no significant interaction between test type and learning condition, Wilks' Lambda = .707, F (2.000, 14.000) = 2.905, p = .088, partial ηp^2 = .293. A one-way ANOVA for test scores averaging across all three test types revealed no overall effect of learning condition (song, prose), F (1, 15) = 2.890, p = .110, ηp^2 = .162.

5.6.1.3 Belgium

Only two experimental groups, a song group and a poem group were available in Belgium, and the number of participants was fairly low. Test results for the free-recall test of the poem group are unavailable, due to an error during data collection. Table 5.5 shows the descriptive statistics for their performance on the three – or two – memory tests in immediate posttesting. Results are provided as a ratio of the total possible scores. In addition, raw mean scores are displayed in brackets.

Table 5.5: Belgium – performance on three memory tests during immediate testing across two learning conditions (song, poem)

	Free Recall (total: 220/225)				Cued Recal (total: 111)	F	Recognitio (total: 18)	n
	N	М	SD	Ν	М	SD	N	М	SD
SONG		.2404	.1028	9	.3774	.1584	9	.5123	.1638
5010	9	(52.89)			(41.89)			(9.22)	
DOFM				15	.3021	.1123	15	.5481	.1326
POLIVI					(33.53)			(9.87)	

At the Belgian institution during immediate posttesting, the song group produced higher mean scores on the cued-recall test, while the poem group yielded the higher mean score on the recognition test. A mixed within-between subjects repeated-measures ANOVA revealed the following: As expected, the factor "test type" was significant, with Wilks' Lambda = .290, F (1.000, 22.000) = 53.744, p = .000, ηp^2 = .710. There was also a significant interaction between test type and learning condition, Wilks' Lambda = .828, F (1.000, 22.000) = 4.568, p = .044, ηp^2 = .172. A one-way ANOVA exploring differences between test scores averaging across two test types showed no significant overall effect of learning condition (song, poem), F (1, 22) = .144, p = .708, ηp^2 = .006.

In delayed posttesting in Belgium, again only data from the learning conditions song and poem were available. Table 5.6 shows the descriptive statistics for their performance on the three memory tests. Results are provided as a ratio of the total possible scores. Again, raw mean scores are displayed in brackets.

	Free Recall			(Cued Recall			Recognition			
	(total: 220/225)				(total: 111)	(total: 18)				
	Ν	М	SD	Ν	М	SD	Ν	М	SD		
SONG	8	.2023	.0816	8	.3300	.1280	8	.5000	.1680		
50110		(44.51)			(36.63)			(9.00)			
POFM	14	.1422	.0366	14	.2915	.0918	14	.5198	.1124		
I OLIVI		(31.28)			(32.36)			(9.36)			

Table 5.6: Belgium – performance on three memory tests during delayed testing across two learning conditions

In delayed posttesting in Belgium, the song group produced higher mean scores in free recall and cued recall. The poem group produced a slightly higher mean score on the recognition test. A mixed within-between subjects repeated-measures ANOVA revealed that the factor "test-type" was significant, Wilks' Lambda = .094, F (2.000, 19.000) = 91.702, p = .000, ηp^2 = .906. There was no significant interaction between test type and learning condition, Wilks' Lambda = .886, F (2.000, 19.000) = 1.218, p = .318, ηp^2 = .114. A one-way ANOVA for test scores averaging across all three test-types showed no overall effect of learning condition (song, poem), F (1, 20) = .448, p = .511, ηp^2 = .022.

5.6.2 Discussion of the overall results at three institutions

I will briefly discuss the overall findings for each institution. A further discussion of all findings produced in this study as well as a reflection on the results in light of the existing research can be found in section 5.7. When looking at the overall results at the three institutions, the first thing to note is that participants display verbatim text retention following all three learning conditions and without prior warning that word-for-word memory would be tested. Furthermore, verbatim retention occurred at all three institutions involved, implying a memory for surface structure in language learners of different proficiency levels, that is, advanced learners (Serbia) as well as high-intermediate learners (Belgium) and low-intermediate learners (Germany). Individual participants from all three learning conditions at the three institutions recalled between 9 and 127 words during immediate free recall and between 5 and 143 words during delayed free recall (with a total possible score of 220 under the song and poem condition and 225 under the prose condition). During cued recall (total possible score: 111),

individual participants recalled between 9 and 103 words on the immediate posttest and between 4 and 96 words during the delayed posttest. Individual results for the recognition test (total possible score: 18) varied from 3 to 16 correctly selected target phrases during immediate posttesting and 4 to 17 correct answers during delayed posttesting. Advanced learners from Serbia produced the highest results on all three memory tests both during immediate and delayed posttesting, while low-intermediate learners from Germany produced the lowest results.

These findings provide evidence that foreign language learners display memory for verbatim text regardless of their proficiency level and regardless of the learning condition they participate in. However, the number of words recalled, whether during free or cued recall, and the number of phrases correctly recognised were influenced by the participants' proficiency level with more advanced learners overall showing greater verbatim retention. It should be highlighted that participants were unaware that they would be tested on their word-for-word memory of the target text. While they were warned that the intervention would be followed by a series of tests, they did not receive an explanation regarding the tests' focus or format. In fact, many participants at the three institutions expressed surprise when they realised that they would be required to recall the original wording.

More importantly, the test results also demonstrate an effect of learning condition on verbatim text retention. Overall, the three intervention studies support the following order of effectiveness: The song condition resulted in the highest verbatim retention, followed by the poem condition. The prose format is the least effective of the three learning conditions in terms of memory for the exact surface structure of the target text. These claims are, above all, supported by test results produced at the largest institution, that is, by advanced learners in Serbia. A comparison of the results produced by the three experimental groups during immediate posttesting in Serbia showed that the song group had the highest mean scores on all three tests. The prose group produced the lowest mean scores on all three tests revealed a significant effect of learning condition. However, post hoc tests revealed a statistical significance only when comparing the song condition and the prose condition. The poem condition produced an intermediate mean score that did not differ significantly from either the song or the

prose group. Yet, during delayed posttesting the poem group again produced higher mean results on all three tests than the prose group, and this difference did, indeed, reach statistical significance. Unfortunately, no data for the song group was available during delayed posttesting. It can only be speculated that the gap in performance might have widened, similarly to the difference between the poem and the prose group. Findings from Serbia, thus, support the claim that working with a song has a positive memory effect compared to working with a prose text. Interestingly, working with a poem format can also have a positive mnemonic effect. Serbian data does not provide conclusive evidence whether the song condition has any advantage over the poem condition.

Findings from the two smaller institutions – Germany and Belgium – show a less straightforward picture. We recall that at both institutions only two experimental groups were available for comparison. In addition, the number of participants was small. Descriptive statistics for the German institution showed that the song group produced higher average scores in free recall and recognition during immediate posttesting. For cued recall, the prose group reached a slightly higher mean score. During delayed posttesting, however, the song group consistently outperformed the prose group on all three memory tests. Still, the overall difference in performance on the three tests during both immediate and delayed testing was not statistically significant. Two conclusions can be drawn from these findings: While the data collected in Germany failed to provide support for the hypothesis that a song condition has a superior mnemonic effect compared to a prose condition, it also does not contradict this assumption. In fact, while differences between the groups' test scores do not reach statistical significance, a large effect size at delayed posttesting implies that learning condition might be a relevant factor and that a greater number of participants might have resulted in significant differences in performance. While this remains a speculative assumption, in the following section, I will explore further whether greater differences are detectable when considering each memory test separately.

Findings in Belgium at first sight do not seem to follow a clear pattern: During immediate posttesting, the song group produced higher average scores on the cued-recall test, and the poem group scored higher on the recognition test. (Data for free recall is missing.)

During delayed posttesting, the song group outperformed the poem group during both free recall and cued recall, but again scored lower on the recognition test. Inferential statistics revealed that these differences were statistically insignificant. Effect sizes were small. These findings indicate that a song format and a poem format might be similar in terms of their effect on learners' memory for verbatim text. Overall, results from these three institutions demonstrate that presenting a text as a song can have a greater mnemonic effect than presenting it in prose format. Therefore, it can be assumed that songs can help learners memorise text and consequently aid language learning. In addition, these findings imply that presenting a text as a poem might be similarly effective as the presentation in the format of a song. In the next section, I will report on and discuss differences between the effects of learning condition on the participants' performance on each separate memory test.

5.6.3 Results for each memory test (free recall, cued recall and recognition)

An additional analysis was conducted, investigating the possible effect of the learning condition on the scores of each test type separately. This was done by means of a T-Test or – in case of a non-normal distribution of scores – its non-parametric alternative, a Mann-Whitney U Test.

5.6.3.1 Free Recall

Free recall is considered the cognitively most demanding of the three tests (see section 5.3.4). Descriptive statistics show that the song condition resulted in the highest average score during immediate and delayed posttesting at all three institutions. Table 5.7 provides an overview of the average scores (presented as a ratio of the total possible scores) gained by all experimental groups in Serbia, Germany and Belgium during immediate and delayed testing.

Institution	Delay	Song N	Song Mean	Song SD	Poem N	Poem Mean	Poem SD	Prose N	Prose Mean	Prose SD
Serbia	imm	13	.3224	.0919	18	.2745	.1044	32	.2739	.1037
Serbia	del	n/a	n/a	n/a	16	.2895	.1297	29	.2547	.1268
Germany	imm	8	.1676	.0708	n/a	n/a	n/a	10	.1178	.0578
Germany	del	8	.1557	.1055	n/a	n/a	n/a	9	.1052	.0514
Belgium	imm	9	.2404	.1028	n/a*	n/a*	n/a*	n/a	n/a	n/a
Belgium	del	8	.2023	.0816	15	.1427	.0354	n/a	n/a	n/a

Table 5.7: Free Recall: Descriptive statistics for Serbia, Germany and Belgium

*Immediate data from Belgium is not available due to problems during data processing

A series of T-Tests, or, if appropriate, Mann-Whitney U Tests (MWU), involving pairwise comparisons of test scores of each experimental group (song – prose, song – poem, poem – prose) at each institution, revealed no significant effect of learning condition. Table 5.8 provides detailed information of the referential statistics, including the tstatistic (or z-statistic), the p-value and the effect size.

Institution	Delay	Comparison	Analysis	р	T-/Z-Statistic	Cohen's d
Serbia	imm	song – prose	TTEST	.15	1.466	0.49
Serbia	imm	song – poem	MWU	.068	-1.822	0.49
Serbia	imm	poem – prose	MWU	.77	283	0.0058
Serbia	del	poem – prose	TTEST	.387	.874	0.27
Germany	imm	song – prose	TTEST	.119	1.646	1.04
Germany	del	song – prose	TTEST*	.247	1.231	0.61
Belgium	imm	imm song – poem		n/a	n/a	n/a
Belgium	del	song – poem	TTEST*	.083	1.969	0.95

Table 5.8: Free recall: Inferential statistics for Serbia, Germany and Belgium

*equal variances not assumed

5.6.3.2 Cued Recall

Cued recall is considered to be cognitively less demanding than free recall as it provides retrieval cues, yet it is assumed to be more demanding than recognition (see section 5.3.4). With one exception, descriptive statistics show that the song groups produced the highest average scores in cued recall at their respective institution. In immediate testing in Germany, however, the prose group outperformed the song group. Table 5.9 provides an overview of the average scores (presented as a ratio of the total possible score) gained on the cued-recall test by each experimental group at each institution during immediate and delayed posttesting.

Institution	Delay	Song	Song	Song	Poem	Poem	Poem	Prose	Prose	Prose
	_	N	Mean	SD	Ν	Mean	SD	N	Mean	SD
Serbia	imm	13	.5343	.1797	18	.4710	.1353	29	.4272	.1354
Serbia	del	n/a	n/a	n/a	16	.5434	.1462	29	.4691	.1554
Germany	imm	8	.2207	.1009	n/a	n/a	n/a	9	.2446	.1051
Germany	del	8	.2444	.1395	n/a	n/a	n/a	9	.1832	.1226
Belgium	imm	9	.3774	.1584	15	.3021	.1123	n/a	n/a	n/a
Belgium	del	8	.33	.128	15	.2883	.0893	n/a	n/a	n/a

Table 5.9: Cued recall: Descriptive statistics for Serbia, Germany and Belgium

A series of pairwise comparisons of mean scores produced by the experimental groups during cued recall revealed a significant effect of learning condition only when comparing results from the song and prose group in Serbia in immediate posttesting. Unfortunately, no data was available for the song group in delayed posttesting in Serbia. Table 5.10 provides detailed information of the referential statistics, including t-statistic (or zstatistic), p-value and effect size.

Institution	Delay	Comparison	Analysis	р	T-/Z-Statistic	Cohen's d
Serbia	imm	song – prose	TTEST	0.039	2.139	0.67
Serbia	imm	song – poem	TTEST	0.271	1.121	0.39
Serbia	imm	poem – prose	TTEST	0.286	1.079	0.32
Serbia	del	poem – prose	TTEST	0.125	1.566	0.49
Germany	imm	song – prose	TTEST	0.64	-0.477	0.23
Germany	del	song – prose	TTEST	0.351	0.962	0.46
Belgium	imm song – poem		TTEST	0.187	1.363	0.55
Belgium	del	song – poem	MWU	0.538	-0.615	0.38

Table 5.10: Cued recall: Inferential statistics for Serbia, Germany and Belgium

*equal variances not assumed

5.6.3.3 Recognition

Recognition is considered to be the cognitively least demanding of the three memory tests (see section 5.3.4). Descriptive statistics show that the song condition resulted in the highest average score at two of the three institutions, that is, in Serbia and in Germany. In addition, the poem group in Serbia outperformed the prose group in both immediate and delayed posttesting. When comparing the scores of the song group and poem group in Belgium, we find that the poem group produced higher mean scores on the recognition test during both immediate and delayed posttesting. Table 5.11 provides an overview of the average scores (presented as a ratio of the total possible score) on the recognition test produced by all experimental groups at their respective institutions during immediate and delayed posttesting.

Institution	Delay	Song N	Song Mean	Song SD	Poem N	Poem Mean	Poem SD	Prose N	Prose Mean	Prose SD
Serbia	imm	13	.7650	.0939	18	.6944	.1115	32	.6458	.1005
Serbia	del	n/a	n/a	n/a	16	.7569	.1297	28	.6468	.1248
Germany	imm	8	.5903	.1453	n/a	n/a	n/a	10	.4611	.2029
Germany	del	8	.5625	.1498	n/a	n/a	n/a	9	.3951	.1533
Belgium	imm	9	.5123	.1638	15	.5481	.1326	n/a	n/a	n/a
Belgium	del	8	.5000	.1680	14	.5198	.1124	n/a	n/a	n/a

Table 5.11: Recognition: Descriptive statistics for Serbia, Germany and Belgium

A series of T-Tests (and a Mann-Whitney U Test) involving pairwise comparisons of the learning conditions revealed a significant effect of learning condition when comparing the results produced in the song and prose condition during immediate posttesting in Serbia. Also, the poem condition yielded significantly higher results than the prose condition during delayed posttesting in Serbia. Furthermore, in Germany the learning condition had a significant effect during delayed posttesting with the song group producing a significantly higher mean score than the prose group. Table 5.12 provides detailed information of the referential statistics, including t-statistic (or z-statistic), pvalue and effect size.

Delay	Institution	Comparison	Analysis	р	T-/Z- Statistic	Cohen's d
immediate	Serbia	song-prose	TTEST	.002	-3.174	1.22
immediate	Serbia	song-poem	TTEST	.076	-1.774	.68
immediate	Serbia	poem-prose	TTEST	.175	-1.355	.46
delayed	Serbia	poem-prose	TTEST	.008	2.775	.86
immediate	Germany	song-prose	TTEST	.15	1.513	.73
delayed	Germany	song-prose	TTEST	.038	2.272	1.11
immediate	Belgium	song-poem	TTEST	.563	587	.24
delayed	Belgium	song-poem	MWU	.743	333	.14

Table 5.12: Recognition: Inferential statistics for Serbia, Germany and Belgium
5.6.4 Discussion of the separate results for each memory test

The results of a series of pairwise comparisons of learning conditions for each of the three memory tests separately revealed the following: Strong differences between learning conditions occur mostly on the cognitively least demanding test type, that is, on the recognition test. In Serbia, the song group and the prose group produced significantly different scores on the recognition test during immediate posttesting. Unfortunately, no data was available for the song group during delayed posttesting. One can only speculate that the difference between these groups on the delayed recognition test might have been at least as strong as the difference that occurred during immediate posttesting. The gap in the average test performance might have also increased, given the finding that the difference between the poem and the prose group increased and reached statistical significance during delayed posttesting. Furthermore, the song group in Germany significantly outperformed the prose group on the delayed recognition test. Large effect sizes in Germany for the recognition tests during both delayed and immediate testing imply that statistical significance for learning condition might have been reached with a larger number of participants. In addition to the strong differences in performance on the recognition test at two institutions, the song group in Serbia also significantly outperformed the prose group during immediate cued recall. No significant differences can be found for free recall at any of the three institutions.

5.7 Overall discussion of the findings of the intervention study

The present quasi-experimental intervention study examined the effects of a song, a poem and a prose text on verbal learning in the context of a typical language lesson. Three trials at secondary and tertiary institutions in Serbia, Germany and Belgium explored the following research question within an ecologically valid classroom setting: Does the presentation of a text in the format of a song result in higher verbatim text retention compared to a presentation in the format of a poem or prose text? In this overall discussion I will interpret the study findings in the wider context of the existing empirical research on verbatim text retention and the use of songs for vocabulary and text learning in a foreign language. I will explore the results step by step, that is, I will discuss relevant findings pertaining to all learning conditions (song, poem, prose) and then move on to the differences between them. Additionally, I will point out the effect of

two variables moderating the superior mnemonic effect of songs, that is, the proficiency levels of learners and the type of memory most strongly affected. Finally, I will discuss the research design in relation to other studies focused on L2-verbal learning through songs, and I will highlight the high ecological validity of this classroom-based intervention study.

5.7.1 Second language learners' retention of verbatim text

Before I discuss the results of the present study in response to the research question, I want to address a more general or basic observation made during the quasi-experiment. The study results give strong evidence that foreign language learners form traces of verbatim text memory. More specifically, the three trials in this quasi-experimental study have demonstrated that verbatim retention of target text occurs in L2-learners following all three learning conditions (song, poem, prose). It seems that L2-learners can display short- and long-term memory of the linguistic form of longer connected text, even if they engaged in a predominantly meaning-focused learning session and even if they were unaware that they would be tested on their word-for-word memory of the target text. This finding contradicts a commonly held belief (e.g. Bartlett, 1932; Jarvella, 1971) that memory for the surface structure of a text disappears quickly and all that remains is the memory of the semantic gist. Instead, this quasi-experiment contributes to the small number of studies that have either directly (e.g. Gurevich et al., 2010) or indirectly (e.g. Wallace, 1994; Tillmann & Dowling, 2007) shown that elements of the surface structure of longer connected text can be retained even long-term and not only by native speakers but also by foreign language learners. The present study showed that such verbatim memory can occur even if learners are primarily focused on meaning and engage in a variety of meaning-focused activities. Furthermore, word-for-word memory was observed in learners of different proficiency levels (low-intermediate, high-intermediate, advanced) and of different backgrounds regarding their L1.

Few studies have explicitly studied the verbatim text memory of second language learners. This study, thus, provides rare evidence that L2-learners display memory for the surface structure of texts. This finding is in line with the findings of a small number of studies that, while not explicitly focusing on verbatim text retention, have used verbatim memory as a measure of verbal learning (Smith Salcedo, 2002; Ludke, 2010). Smith

Salcedo (2002), for example, tested beginning-Spanish learners' verbal learning following a song or spoken intervention by means of a cloze test. While Smith Salcedo used the cloze test as a comprehensive measure of language skills, this test can effectively be considered a cued-recall test of verbatim text retention. Smith Salcedo's participants displayed verbatim memory of longer connected text by showing cued recall of words from all parts of the target lyrics. Ludke (2010) focused on the acquisition of short expressions in an unknown language and measured the cued recall of short pairedassociate phrases in Hungarian and English by English speakers without any prior knowledge of the Hungarian language. The participants in this case were clearly aware of the fact that retention of the exact wording of the phrases was the purpose of the study. They were thus "warned" and were also encouraged to orally rehearse the Hungarian phrases and, thus, would have also had a highly form-focused approach to learning of the target phrases while, of course, also trying to remember the meaning of these entirely novel expressions. In Ludke's research all learning conditions – sung, spoken and rhythmical spoken – produced clear signs of correct memory of the target phrases.

5.7.2 The effects of a song-based and a prose-based treatment on verbatim text retention

I now want to address the main research interest of the present quasi-experiment and discuss the effects of songs on verbal learning compared to other text formats. We recall the research question of the current study: Does the presentation of a text in the format of a song result in higher verbatim text retention compared to a presentation in the format of a poem or prose text? The intervention study has demonstrated that songs can result in significantly higher verbatim text retention compared to prose texts. This result corresponds with the results in a number of studies in the field of cognitive psychology, which showed a beneficial effect of a musical presentation on verbatim text memory in the first language (Wallace, 1994; McElhinney & Annett, 1996; Calvert & Tart, 1993; Purnell-Webb & Speelmann, 2008). It also supports findings of a small number of studies in second language research showing that L2-learners can display significantly higher verbal retention of phrases or longer text following a song compared to a spoken text (Hahn, 1972; Ludke, 2010; Smith Salcedo, 2002). More specifically, this outcome corresponds with findings in one of Ludke's (2010) controlled laboratory experiments

which showed that a sung learning condition resulted in greater verbal retention than a spoken and rhythmic spoken condition (see Chapter 2 for a detailed discussion of Ludke's research). However, while Ludke's study revealed such differences primarily on highly demanding productive tests, the current quasi-experiment showed the strongest differences between the learning conditions on the cognitively least demanding tests, that is, the recognition test and, in one instance, the cued-recall test. This divergence from previous research findings will be discussed in greater detail below. The findings also support Smith Salcedo's (2002) findings that listening to songs can result in higher cued verbatim recall (measured by a cloze test) in beginning-Spanish learners compared to a spoken presentation in a natural classroom setting. However, while Schmitt Salcedo could only detect significant differences during immediate posttesting and none during 2-weeks delayed posttesting, the current study demonstrated superior results – albeit for a very small number of participants (in Germany) – during 1-week delayed posttesting as well. It can only be speculated that the song group in Serbia might have also produced retention rates superior to the Serbian prose group.

5.7.3 The effect of a poem-based intervention on verbatim retention

So far, I have focused on the comparison of the song and the prose condition. However, the current study also involved a poem condition. The test results overall located the poem group between the other two learning conditions. Descriptive statistics indicated that the poem-based intervention resulted in lower mean scores than the song-based treatment in Serbia during immediate posttesting on all three memory tests. In Belgium, a similar pattern emerged on the immediate cued recall test and for delayed free and cued recall. However, results were reversed for the recognition test in Belgium, with the poem group producing the higher mean score during both immediate and delayed posttesting. Regardless, a pairwise analysis revealed that the differences between the song groups and the poem groups at these two institutions had no statistical significance. The poem-prose comparison, on the other hand, produced a clearer pattern in the descriptive data. In Serbia, the only institution that allowed for a direct comparison of the poem and the prose condition, the poem group produced higher mean scores on all tests during both immediate and delayed posttesting. This difference was statistically significant overall for the delayed results. A separate analysis of test performance on

each of the memory tests separately revealed that this difference was only significant, and then highly so, on the recognition test (p = .008). This is in accordance with the overall finding that differences between learning conditions were predominantly detectable on the less demanding memory tests.

The results regarding the effect of a poem-based teaching unit reveal that on the one hand presenting a text as a poem can result in superior test performance compared to a prose format in terms of verbatim text retention. On the other hand, a song-based treatment does not seem to be significantly more effective than a poem-based treatment, at least in regard to long-term memory. In other words, a poem condition can be similarly effective as a song condition. This finding provides some evidence that the melody, while probably playing a part in the mnemonic effect of songs, is not the only factor rendering song lyrics more memorable than prose texts. Instead, particular characteristics of the lyrics themselves seem to render song lyrics but also poems more memorable. This result is in accordance with Tillmann and Dowling's (2007) finding that L1-speakers had significantly greater recognition memory for surface details of a poetic text with rhythm and typical structural regularities such as rhyme, compared to a prose text. In Chapter 6, I will discuss in detail what factors might be decisive for the memory effect of songs and poems. In addition, I will discuss pedagogical implications and address the question whether poems could and should be used in the classroom instead of songs.

It has to be acknowledged that the defining characteristics of the poem, particularly in contrast to the prose format, in this research design remain somewhat vague. Mainly, this format differed from the prose text in that it was encouraged to be perceived as a poem by the participants. This was due to the fact that it was explicitly introduced to the subjects as a poem while the prose text was introduced as a funeral speech from a movie. More importantly however, the poem was visually presented in poetic format, that is, it was printed in a line-by-line structure rather than a running text. It also preserved the metre and rhyme (including internal rhyme such as assonance and alliteration) of the song lyrics. However, to a large extent so did the prose text. Only in few instances was the poetic metre in the prose version interrupted by the addition or deletion of up to three words or the reversal of the main and subordinate clause in one sentence. However, the reader, or rather performer of the two text versions was

encouraged to accentuate the rhythm to a greater extent in the poem than in the prose text. While this does – when listening superficially – not seem to make a great difference, it is possible that these differences might have affected the learners' comprehension, processing, encoding and later retrieval of the information.

5.7.4 Moderating variables: the participants' language proficiency and the test type

I will now take a closer look at other variables moderating the effect of the learning condition. Firstly, I will discuss the relevance of the learners' proficiency level for verbatim text retention and verbal learning through songs. While many studies have focused on beginning learners (e.g. Hahn, 1972; Medina, 1990; Smith Salcedo, 2002) or even learners without any prior knowledge (Ludke, 2010) of the target language, the quasi-experiment at hand looked at learners with varying proficiency levels, that is at advanced, low-intermediate and high-intermediate learners. The three trials of the quasiexperiment showed that verbatim text retention occurs in learners of all three proficiency levels, though to differing degrees. As expected, advanced learners showed the greatest retention and low-intermediate learners displayed the lowest rate of remembered words, while the high-intermediate learners were located between these two. These results demonstrate that memory for verbatim text occurs for both intermediate and advanced learners, and they provide support for the use of songs with learners of all of these levels to teach vocabulary based on the memorization of connected text. However, as I will elaborate further in the next paragraph, the incidental or intentional - memorization of phrases and longer connected text can be demanding. As the study demonstrated, the lower the learners' proficiency level, the lower the number of words they remembered.

This brings me to the cognitive demand of the task: As discussed previously, a significant difference between learning conditions was predominantly observed on the cognitively least demanding task, that is, the recognition test. Recognition can be considered cognitively least demanding, while free recall was the most challenging of three memory tests used in the present study (cf. Tulving, 1985). This finding is in clear contrast to Ludke's (2010) predictions (based on Yalch, 1991) and experimental findings, that differences in test performance between a sung and a spoken learning condition were observed primarily on the most demanding tasks involving clear productive elements.

The divergent result of the present study might be explained by the long text used as target material. Ludke, in contrast, focused on the memorization of short unconnected phrases. In addition, it can be assumed that the unannounced nature of the verbatim memory test in the current study rendered the recall and recognition tasks particularly demanding. Effectively, any verbatim retention in the present study was the result of incidental memorization, while Ludke's participants had been explicitly told that the purpose of the experiment was the verbatim memorization of paired-associate target phrases. We recall that the treatment used under all three learning conditions in the study at hand was intended to reflect actual classroom practices as much as possible. It therefore had focused on meaning and comprehension and involved activities such as comprehension questions and a personalised discussion. At the same time, the treatment also comprised form-focused activities such as the gap-fill and the sing- or read-along. Still, the memorization of text was not an explicit goal, and while the participants were aware that a test battery would follow, they were not made aware that the tests would target verbatim text memory.

This finding indicates that overall verbatim memory of such a long connected text after only three repetitions and in a predominantly meaning-focused lesson (and particularly without prior warning that verbatim retention was the aim of the task) is highly demanding, particularly for lower-level learners. If memorization of some form should be, in fact, the purpose of a classroom activity, a more explicit form-focus might be advisable. Barcroft's (2002) types of processing-resource allocation (TOPRA) model proposes that processing capacities for form and meaning are in competition and are depending on the cognitive demand and therefore also on the proficiency level of the learners - allocated to one processing type at the expense of the other on first encounter. In other words: If learners struggle to understand the meaning because the complexity of the text and possibly the additional musical accompaniment result in a highly demanding task in terms of listening comprehension, vocabulary, grammar or simply the overall length of the text, then learners might not be able to "spare" much cognitive capacity for an elaborate processing and for the retention of surface structures. Anecdotal evidence sometimes indicates the opposite, since many teachers can point out students who have an impressive memory of song lyrics that are beyond their proficiency

level. In addition, the way teachers frequently implement song-based lessons, including the repeated listening and the engagement in both form- and meaning-focused activities, might free up cognitive resources and permit successive or increasingly simultaneous processing of both meaning and form. Further pedagogical implications derived from these results for the use of songs with learners of varying language proficiency will be discussed in Chapter 6.

5.7.5 The ecological validity of the intervention

Finally, one defining element of the present study design needs to be discussed within the context of the existing research: The study at hand strove for a high degree of ecological validity. The pronounced purpose was not to conduct a highly controlled laboratory experiment. Equally, it was at this point not of concern to investigate the effect of an "ideal" classroom intervention informed by cognitive-psychological research. Instead, the current study was intended to lay some groundwork and, thus, focused on verbal learning with the help of songs as they currently tend to be used in language classrooms around the world. The treatment lesson was therefore closely informed by the results of the teacher questionnaire discussed in Chapter 3 regarding, among other things, the length of the intervention, the choice of activities, the number of repetitions and the choice of treatment song. The survey respondents' decisions were influenced by their own beliefs as well as curricular and time-constraints. The current study provides evidence that songs as they are currently used in the language classroom can result in verbal learning or more specifically, in word and text memorization. This study, thus, adds to the currently existing body of research consisting predominantly of studies which have used a more controlled design in a laboratory setting (Ludke, 2010) or which conducted classroom-based research that in many respects did not entirely reflect current teaching practice (Hahn, 1972; Smith Salcedo, 2002; Salas, 2006; Ludke, 2010).

Particularly the time investment before and during the teaching unit often required by the experimental design in some of these studies seem to be not entirely reflective of actual teaching practices under everyday curricular and time constraints. This claim is supported by the teacher survey (Study 1). Teachers responding to the questionnaire reported replaying a song an average of two to three times during one lesson. In addition, most informants only sometimes replayed a song during a subsequent lesson.

Some studies also involved additional and somewhat innovative methods. Hahn (1972) made extensive use of pattern drills characteristic for the audiolingual teaching approach of the time but less commonly found in the currently favoured communicative language teaching (Richards, 2005). Medina (1990) introduced an additional independent variable into her study in the form of illustrations. Schmitt Salcedo's (2002) treatment involved six repetitions of one song over several lessons. Ludke's (2010) quasi-experimental classroom-based study involved six 15-minute lessons focused on only one song over the course of two weeks and involved fairly elaborate activities such as a video-recording of the class performance of the song and the subsequent production of a DVD. Salas (2006) is a professional singer. She used two songs in her study that she herself carefully composed for the purpose. In addition, she used a very particular method called the "audiation technique", an intentional subvocal rehearsal method usually employed by singers to memorise lyrics. These studies provided valuable evidence that singing and songs can have a mnemonic effect and can result in higher verbal learning compared to spoken and also rhythmic spoken text presentations – however, in controlled laboratory settings or somewhat "idealised" classroom settings. The current study provides an additional angle and demonstrates that songs as they currently tend to be used can benefit language learning.

5.8 Limitations of the intervention study

This study has a number of limitations, some of which might have a fairly strong impact on the results and consequently the findings of this quasi-experiment. The more obvious limitations which have already been discussed previously are the lack of delayed data from the song group in Serbia and the missing free-recall test results from the poem group in Belgium. In addition, the inability to compare all three learning conditions at the smaller institutions in Belgium and Germany resulted in the comparison of only two – and two different – groups at each of these institutions. This rendered the comparison between institutions rather more complicated and prevented a collective statistical analysis. It has also already been mentioned that the number of participants was small. The smallest groups were found in Germany with only eight participants in the song group and nine participants in the prose group. Even at the largest institution, Serbia, the poem group consisted of only 14 participants. In addition, in Serbia the group sizes varied

considerably from 14 in the poem group to 32 in the prose group. The frequently small size of experimental groups as well as the great differences between groups might have affected test results. Small group sizes might have resulted in individual learner differences having a greater influence. In fact, when analysing the results for each memory test separately, the data in at least one group was not normally distributed. This limitation highlights the fact that potential individual differences between learners were not controlled for. Ludke (2010), for example, controlled for a number of individual differences and explored their impact on the effect of the learning condition. And although many of these learner characteristics did not influence test performance, Ludke found that prior language learning experience, as well as a positive mood at the beginning of the experiment and participants' music production abilities significantly affected test results.

It also needs to be acknowledged that the ecological validity of the intervention was informed by the teacher survey (Chapter 3) and reflected typical teaching practices involved in song-based teaching. However, the way the comparison treatments, that is, the poem- and prose-based interventions, were designed was not informed by such a survey, and it cannot be guaranteed that the ways the students in these conditions dealt with the text were also "typical". However, as detailed in the methodology section above, care was taken that no teaching practices were used that can be considered inappropriate for these other two text formats. Another ground for caution is uncertainty about the participants' prior knowledge of the target text. It is not inconceivable that the song lyrics were vaguely familiar to some of the participants. On the other hand, it would seem unlikely that all participants with some prior knowledge of the text happened to find themselves in the song groups. Nevertheless, it needs to be debated whether some participants might have been somewhat familiar with the target song, which might have given an advantage to the song group, as music could have functioned as a memory trigger cueing the recollection of the previously heard lyrics. However, the poem group participants were not exposed to the melody during the intervention and were not alerted to the fact that their target poem was based on song lyrics, thus limiting the effects of possible memory cues. Still, the poem-based intervention overall resulted in significantly higher verbatim text memory compared to the prose-based treatment.

A further limitation is the use of only one song, as the song's particular characteristics might have had different effects for the different learning conditions. Using more than one target song would have permitted a more thorough investigation into musical and lyrical characteristics and their effects on verbatim text retention. In fact, contrary to Wallace's (1994) recommendations of a simple and symmetrical melody, in the present study the song's melody was in fact fairly complex with five lines in a verse, rendering the verses melodically unsymmetrical. In addition, the lyrics frequently did not display clear end-rhymes and instead frequently displayed semi-rhyme or line-final assonance. However, these factors, rather than giving the music group an advantage in terms of verbatim text retention, should have rendered the song condition less likely to result in superior test performance. Unfortunately, quasi-experimental research involving more than one song was practically not feasible in this project. The institutions that lent their participation to the research project could not spare more time and let their students participate in research that did not form part of the official curriculum.

This problem also eliminated initial plans of a counter-balanced study design that would have allowed for all participants to engage in all learning conditions and would thus have controlled for potential individual differences. In addition, a counter-balanced approach would have changed the basic premise to investigate verbatim text retention following a meaning-focused lesson and without prior warning of the type of learning required. It can further be criticised that the participants' proficiency level was not specifically assessed prior to the experiment. Instead, the researcher relied on the teachers' and/or institutions' assessment of the classes overall performance level. This approach neglects individual differences between participants in each group. Furthermore, this approach lacks a reproducible assessment of proficiency based on the same criteria for all three institutions.

5.9 Conclusion

In summary, the main findings of the quasi-experimental intervention study provided evidence that the presentation of text in the format of a song can result in higher retention of verbatim text both in immediate and one-week delayed testing. Such a beneficial effect can be observed in learners of both intermediate and advanced

language proficiency. Additionally, presentation of text in the format of a poem can also result in higher retention compared to a prose text. Greater verbal retention in the song groups and the poem groups (compared to the prose groups) was mostly observed on the least demanding task, that is, the recognition test. These results give some indication that the benefits of a song format or poem format might not be pronounced enough to be reflected in the performance on more demanding tests. In the following chapter – the overall discussion – I will discuss possible reasons why songs and poems as used during the intervention might benefit verbatim text retention. Furthermore, I will discuss pedagogical implications of the study.

Chapter 6 Overall discussion

6.0 Introduction

This chapter discusses the main results of the three studies presented in this thesis. The teacher questionnaire (Study 1), the song-corpus study (Study 2) and the quasi-experimental intervention study (Study 3) addressed different research questions. However, like puzzle pieces, their findings ultimately produced one overall picture and provided a coherent answer to one overarching research interest. In this overall discussion, I will first briefly repeat the main results of each study separately before presenting the overall findings of the thesis in response to the primary research question: Can songs as they are currently used in foreign and second language classrooms benefit lexical learning? Furthermore, this chapter explores possible factors underlying a mnemonic effect of songs and poems as well as pedagogical implications that result from the findings of this research project. I will also highlight the original contributions of this thesis to the research on songs in the foreign language classroom. Finally, I will discuss possibilities for future research and the limitations of this thesis.

6.1 Study 1: The teacher survey

The online questionnaire in Study 1 explored teachers' cognitions regarding the use of songs in the language classroom and their teaching practices involving songs. More specifically, it investigated whether second and foreign language teachers used songs and if yes, how and for what purposes. Furthermore, it examined internal and external factors influencing teachers' choices regarding a musical teaching approach, including conflicts with the official curriculum, the availability of suitable songs and the reaction of the students. Finally, the study explored the use of songs particularly to teach vocabulary. It investigated whether informants employed songs to introduce new words or to consolidate or deepen knowledge of (half-)familiar words. It also inquired whether respondents utilised songs to teach multi-word units, such as phrasal verbs and idiomatic expressions. Additionally, informants were asked about aspects that, in their eyes, rendered songs suitable for vocabulary learning. A total of 568 teachers working in 41 countries and teaching 20 different languages participated in the study. The wide scope of informants in terms of the languages they taught and the countries they worked in

sets this study apart from previous research, which focused on English instructors only and was limited to a small sample of teachers in either the USA (Edwards, 1997; Bjorklund, 2006) or Spain (Pérez Aldeguer and Leganés Lavall, 2012). The results of the study revealed that many teachers display a positive attitude towards songs and actually use them as teaching material in class. Rather than viewing them as special treats and time fillers without a specific language-pedagogical purpose, they frequently utilise them with clear meaning- and language-focused goals in mind and in the context of a directed and diverse teaching unit that can include a variety of complementary activities. One major purpose is the teaching of vocabulary. The present survey revealed that many language teachers use songs to support the acquisition of words and multi-word units, both by means of introducing new words but also by consolidating and expanding knowledge of already familiar words. This relevant finding informed the two subsequent studies, which focused on the vocabulary load and the vocabulary learning potential of teacher-selected songs (Study 2) and the retention of words and longer connected text following a song-based lesson (Study 3).

6.2 Study 2: The pedagogical song corpus

The pedagogical song corpus consisted of 177,384 tokens from 635 songs derived from English textbooks, English-teacher websites and recommendations made by the survey informants. It was compiled with a view to investigate the typical characteristics of songs used in the ESL or EFL classroom. In particular, this corpus study investigated the lexical load of teacher-selected songs, both in regard to their lexical demand as well as their potential for lexical learning, and compared it to the vocabulary demand of other authentic text genres. The main findings of the corpus study can be summarised as follows: The lyrics of songs used by teachers in the foreign or second language classroom tend to be short and repetitive and they tend to have a low vocabulary load compared to other authentic text genres. Consequently, they can be used as "entry-level" authentic and unsimplified texts for non-native speakers. This finding corresponds with the survey informants' highly rated purpose of teaching authentic language and culture through songs. However, the results indicate that teacher-selected songs might still be fairly challenging in terms of lexical demand for lower-level learners. As the previous survey study showed, teachers appear to take the lexical challenge of the lyrics into

consideration, as they frequently assist their students in the comprehension of the lyrics and their vocabulary with a variety of activities. In contrast, for higher-level learners teacher-selected songs seem to offer limited opportunities in terms of learning new vocabulary beyond the most frequent 4,000 English word families, including mid- and low-frequency words and academic vocabulary. The predominant use of high-frequency words, the presentation of vocabulary in context and the high repetition of words in the lyrics suggest that songs, particularly when used with advanced learners, can be more useful to consolidate already familiar words and to support the acquisition of deep word knowledge beyond a first form-meaning connection, including grammatical functions, collocations and associations. This finding is in accordance with teaching practices reported by teachers in Study 1, as a majority of informants reported using songs to consolidate knowledge of familiar vocabulary.

6.3 Study 3: The quasi-experimental intervention study

Study 3 examined the effects of songs on verbal learning in the context of a typical language lesson. Three trials of a quasi-experimental intervention study at secondary and tertiary institutions in Serbia, Germany and Belgium explored the following research question within an ecologically valid classroom setting: Does the presentation of a text in the format of a song result in higher verbatim text retention compared to a presentation in the format of a poem or prose text? The treatment implemented during the quasiexperiment was designed to reflect actual teaching practices as closely as possible. This is different from previous L2-studies which used a controlled laboratory design or employed a classroom-based design which does not bear much resemblance to what my survey (that is, Study 1) revealed to be current common practice. In order to ensure that the activities as well as the choice of material used in the intervention study reflected actual teaching practice as much as possible, the activities and the selection of a song for the quasi-experimental treatment was informed by research findings from the survey study (Study 1) and the song corpus analysis (Study 2). The main findings of Study 3 can be summarised as follows: The presentation of text in the format of a song compared to a prose text can result in higher retention of verbatim text both in immediate and oneweek delayed testing. Such a beneficial effect can be observed in learners of both intermediate and advanced language proficiency. Additionally, the presentation of text in

the format of a poem can also result in higher verbatim retention compared to a prose text. The better verbal memory in the song groups and the poem groups (compared to the prose groups) was mostly observed on the cognitively least demanding task, that is, the recognition test. These results give some indication that the benefits of a song format or poem format might not be pronounced enough to be reflected in the performance on the more demanding recall tests. For a more detailed discussion of the results in consideration of the existing research in the fields of cognitive psychology and second language acquisition see Chapter 5.

6.4 Can songs as they are currently used in second language classrooms benefit lexical learning?

The three studies discussed above established how songs are typically used inside second and foreign language classrooms and showed that such teaching practices can result in increased retention of verbatim text compared to a teaching unit around a prose text. In response to the overall research question this thesis has, thus, provided some evidence that songs, as they are often incorporated in lesson plans around the world, can aid lexical learning, at least as far as retention of verbatim text and linguistic form is concerned. Before I discuss implications for the use of songs in language teaching resulting from this research finding, I first want to discuss possible reasons why a songbut also a poem-based lesson can lead to superior verbatim retention compared to a teaching unit around a prose text. As the design of the study included three different learning conditions (song, poem, prose) which differed in various respects, the study results allow us only to speculate about the potential factors contributing to a mnemonic effect of songs and poems under the present conditions.

6.5 Potential factors contributing to a mnemonic effect of songs and poems

As discussed in the literature review in Chapter 2, various studies in the field of cognitive psychology and in second language acquisition offer a variety of reasons why presenting text as a song (and in few cases also as a poem) can be expected to result in greater verbatim retention of words, phrases and longer connected text compared to prose texts. Interestingly, several of the factors referred to in that literature do not seem to play a major role or at least not the decisive role in the present quasi-experiment. Kilgour

et al. (2000), for example, found that the typically slower speech rate in songs resulted in superior memorization of sung compared to spoken lyrics. However, in the present study all three text formats were presented at a similar speech rate. Multiple repetitions of a song have been considered essential to the mnemonic effect of songs (e.g. Calvert & Tart, 1993) and the repetition of single words, phrases, and lines within one and the same song has also been proposed as a possible reason for the high memorability of lyrics (e.g. Hyman & Rubin, 1990; Foncesa Mora, 2000). Verbatim repetition has been found to play a role in verbal learning (Ebbinghaus, 1885), including the incidental learning of multi-word items (Durrant & Schmitt, 2010). A majority of informants in the survey study concurred with this view that one of the benefits of songs is the possibility to repeatedly listen to and sing along with a song without causing boredom on the part of the students. The potential of using songs as enjoyable drill activities will, thus, be further discussed in the section on pedagogical implications. Still, in this quasiexperiment, all three text formats contained the same words, which were repeated the same number of times, as were the texts as a whole. The better post-test performance in the song (and poem) condition can therefore not be attributed to this repetition as such.

Melody has frequently been the focus of memory research on songs and is understood to have played a crucial role in a number of studies investigating verbatim text retention by native-speaker participants (Calvert & Tart, 1993; Wallace, 1994; McElhinney & Annett, 1996; Rainey & Larsen, 2002). In addition, the melody appears to have been the deciding factor in Ludke's (2010) study on phrase learning in a foreign language. However, in the current study, the close to equivalent posttest performance under the song and the poem conditions suggests that factors other than melody are involved in the mnemonic effect of both formats. While all of the factors referred to in previous research might play a role overall in rendering lyrics more memorable, it seems we should focus on characteristics common to both the song and the poem that most plausibly explain the attested superior verbatim retention as compared to the prose condition.

The following three factors should be considered: Firstly, I propose that the rhythm and the emphasised metre (but also the melody) highlight structural regularities, that is, poetic features, of the lyrics in both the song and the poem and draw the learners' attention to the surface structure of the text. These particular characteristics of songs

and poems might not only draw attention to the formal aspects of the text and encourage processing of form; they might also lead to chunking of words into greater units and thus facilitate retention. As explained in Chapter 2, chunking refers to the binding together of information into larger units and constitutes a way to circumvent the limited processing capacities of the working memory, more specifically, of the phonological loop (Baddeley, 2000). In the following section, I will provide examples of sound patterns found in the lyrics of the treatment song, and I will further discuss possible effects of language-structural regularities on retention.

Secondly, the presentation of the song and the poem on paper in a poetic, that is, in a line-by-line format might have been a factor contributing to higher memorization of verbatim text following the song and poem learning conditions. The line-by-line format might render phonological and rhythmic features of the text more noticeable to the learners. In addition, the visible phrasing of the text might emphasise semantic units, which could again encourage the processing of larger linguistic chunks. In addition, the so-called poetic format could facilitate the segmentation of the text by the readers and might consequently aid reading and listening comprehension and finally also text retention. The effect of the written format is also discussed in greater detail below.

Thirdly, it seems likely that the multimodality of the input experienced during a songbased lesson might result in a particularly rich episodic memory comprising a number of strongly associated retrieval cues such as rhythm, sound patterns, mental imagery, emotions and contextual information from the learning event. A strong episodic memory involving a variety of associated memory traces might not only form a solid basis for the formation of abstracted semantic memory but might also aid the noticing (Schmidt, 1990) of target vocabulary outside the original learning context and thus increase the likelihood of a target item to be learnt. The formation of episodic memory is considered further in section 6.5.4 below.

6.5.1 Structural regularities and a greater attention to surface form

One reason for better verbatim retention of song lyrics and poems might be the high use of poetic devices relying on sound patterns, such as alliteration, assonance and rhyme, and the greater salience of these language-structural features in song lyrics and poems

compared to prose texts due to the rhythm, emphasised metre and also the melody (cf. Wallace & Rubin, 1991; Tillmann & Dowling, 2007). In addition, the way songs are implemented in a second language lesson might also facilitate noticing and processing of linguistic form. Results of the intervention study provide some indication that participants in the song and poem conditions overall showed greater intake of linguistic surface structures. Particularly, the significantly higher performance on the recognition test by the Serbian song group during immediate posttesting and the German song group as well as the Serbian poem group during delayed posttesting (compared to the corresponding prose groups) demonstrated a greater retention of verbatim text and consequently a superior memorization of surface form. As explained in Chapter 5, the recognition test was designed to measure memory for (word-for-word) form rather than for meaning or gist: During the design of the distractors, care was taken that their semantic content did not differ greatly from the original target phrases. Open-class words were replaced by synonyms or near-synonyms (for example, "... leave your <u>worries/doubts/fears/troubles</u> behind ... ") that preserved the gist of the original sentence. In addition, many of the differences between the original phrases and the paraphrases affected function words and consequently did not differ greatly regarding their semantic content (for example, "And would you find the/that one you're dreaming of/about?"). Higher performance on the recognition test was therefore more likely to be a result of form retention than memorization of gist.

Regarding general language use, it has been accepted that the processing and memorization of meaning takes precedence over the processing and retention of form in native speakers (e.g. Jarvella, 1971). Concerning second language acquisition, VanPatten (1989) also argued that the processing of form and meaning of target input in adult second language learners are in competition, and he elaborated that "conscious attention to form in the input competes with conscious attention to meaning and, by extension, only when input is easily understood can learners attend to form as part of the intake process" (p. 414). However, in songs and poems the rhythm, the accentuated metre (and also the melody) might emphasise linguistic form and draw the learners' attention to the surface structure of the text. Such awareness-raising could result in greater cognitive elaboration of linguistic form (cf. Wallace, 1994) and noticing of surface

structure. As (Schmidt, 1990, 2001) argued, second language learning requires noticing, that is, they need to pay attention to relevant features in some form. In addition, song lyrics and poems already frequently contain salient poetic features such as rhyme, alliteration and assonance, which in themselves draw the readers' and listeners' attention to the formal aspects of the text and are also considered to be inherently more memorable than language lacking such noticeable patterns (Wallace & Rubin, 1991). Research in the field of second language acquisition has shown that L2-learners display greater recall of inherently salient sound patterns. Boers and Lindstromberg (2005), for example, demonstrated that EFL learners showed significantly greater recall of alliterative words strings than of non-alliterative word strings, both when studied as decontextualised items (see also Boers, Lindstromberg & Eyckmans, 2013) or under incidental learning conditions when encountered in longer texts while primarily focusing on content (Boers, Lindstromberg & Webb, in press). Such a mnemonic effect was primarily observed in the recall of form and less so in the recall of meaning (Boers, Lindstromberg & Webb, in press). Assonance has also been found to benefit the memorization of multi-word units when encountered in a context that promoted the noticing of form, albeit to a lesser degree than the seemingly more salient sound pattern of alliteration (Lindstromberg & Boers, 2008). Overall, Boers, Lindstromberg and Eyckmans (2014) emphasised that the effect was clearly larger when the participants' attention was explicitly drawn to the sound patterns.

The song lyrics used in the intervention study comprised a high number of sound patterns associated with increased verbatim memorization as exemplified in Figure 6.1. While these structural regularities could be found in all three text formats and thus cannot be seen as the crucial factor contributing to a mnemonic effect of songs and poems, the rhythm and accentuated poetic metre of the treatment song and also the poem seemed to have emphasised at least some of these sound patterns and consequently might have rendered them more noticeable.

STRUCTURAL REGULARITY	EXAMPLE			
Repetition	leave – leaveworth – worth			
Alliteration	 <u>gi</u>ft – <u>gi</u>ven <u>t</u>ry – <u>t</u>ake – <u>t</u>ravelled <u>d</u>onate – <u>d</u>ime <u>f</u>inally – <u>f</u>all 			
Consonance	 fir<u>st st</u>ep – longe<u>st st</u>ride <u>wh</u>at – <u>w</u>orth – al<u>w</u>ays – <u>w</u>orth 			
Assonance	 ag<u>ai</u>nst – gr<u>ai</u>n – w<u>ay</u> life – prize – fight – like – twice – ride – life rewind – life – time – side 			
End rhyme (including assonance)	 last – past above – love advice – right – behind – by – stride 			

Figure 6.1: Examples of structural regularities in the experimental song

Consider the following line from the target text: *Every second counts 'cause there is no second try*. In the song, the "repeated" homonyms *second* (noun) and *second* (adjective) are clearly prosodically emphasised and also slightly prolonged. The same can be said for a number of structural regularities found in the song lyrics. In order to assess whether these (accentuated) structural regularities contributed to the higher memorization of verbatim text under the song and the poem condition, it needs to be analysed whether experimental participants in the song and the poem groups displayed significantly higher retention particularly of these patterns compared to the prose-group participants. Performing a full statistical analysis of the retention of structural regularities by participants under the different learning conditions is outside the scope of the present study. However, a comparison of the immediate free and cued recall of a small sample of words by participants in Serbia provides interesting insight into the retention of words that form part of sound patterns. Table 6.1 shows retention rates of words found in final position of the lines in the written song lyrics. These words displayed end-rhyme or other types of line-final sound patterns such as assonance. In contrast, in the prose format,

these words formed part of the running text. (Only words occurring once in the target text were included in the table and the present analysis.)

	FREE RECALL				CUED RECALL			
	Song	Song	Prose	Prose	Song	Song	Prose	Prose
	(N=13)	(%)	(%)	(N=32)	(N=13)	(%)	(%)	(N=29)
right	7	53.8	34.4	11	13	100	75.9	22
behind	5	38.5	9.4	3				
by	4	30.8	15.6	5				
stride	1	7.7	0	0				
fight	1	7.7	9.4	3	7	53.8	51.7	15
twice	7	53.8	3.1	1	9	69.2	20.7	6
see	4	30.8	25	8				
memories	10	76.9	50	16				
enemies	8	61.5	56.3	18				
above	8	61.5	62.5	20	11	84.6	75.9	22
love	10	76.9	84.4	27	8	61.5	58.6	17
heart	1	7.7	56.3	18	11	84.6	100	29
stars	11	84.6	53.1	17	4	30.8	17.2	5
side	1	7.7	15.6	5	9	69.2	37.9	11

Table 6.1: Retention of words in line-final position in immediate recall in Serbia

As Table 6.1 shows, line-final words displaying end-rhyme (or assonance, semi-rhyme, and oblique rhyme) tend to be better recalled by participants in the song condition compared to the prose condition. A chi-square test for independence (with Yates continuity correction), comparing the proportions of successful and unsuccessful free recall of these fourteen target words in the free recall test, indicates a significant association between learning condition and target word retention, χ^2 (1, n = 630) = 4.07, p = .04, with the song group showing superior free call compared to the prose group. A chi-square test for independence (with Yates continuity correction), comparing the proportions of successful and unsuccessful cued recall of the words in the cued recall test indicates a near-significant association between learning condition between learning condition and target word retention, χ^2 (1, n = 588) = 3.49, p = .062. These findings lend tentative support to the hypothesis that sound patterns are more noticeable and consequently more memorable in the song lyrics and in the poem and that this might, thus, be a factor contributing to the higher retention of songs and poems compared to a prose text. Further research is necessary to explore the effects of different structural regularities on the retention of

words and longer connected texts by means of a detailed items analysis of recalled and recognised target words and phrases.

At this point, I need to address a potential downside of the emphasis of linguistic structure inherent in songs and poems, as it could result in the processing of form at the expense of meaning, at least during the first encounter with the material. According to Barcroft (2002) type of processing-resource allocation model (TOPRA), a second language learner's cognitive processing capacity is limited. Depending on the overall cognitive demand of a task, one type of processing is emphasised at the expense of another when first engaging with verbal material. In the case of foreign-language songs and poems, the salience of structural elements could encourage processing of form at the cost of decreased semantic processing. However, while the results of the intervention study implied that under the song and poem condition the participants' attention is drawn to the surface structure of the target text, the survey study revealed that this awareness of form does not have to come at the expense of meaning. As findings from the teacher survey indicate, informants typically implement a variety of complementary activities that explain and elaborate the meaning of a song, such as comprehension questions, discussions, ordering activities and writing tasks. In addition, the questionnaire respondents reported typically repeating a song two to three times. The repeated encounters with the text are likely to free up cognitive resources as learners get increasingly familiar with the text, permitting successive or (increasingly) simultaneous processing of both meaning and form.

6.5.2 The effect of the written line-by-line format

One apparent difference between the song lyrics and the poem on the one hand and the prose text on the other lies in the written format. All three text types had verses (song, poem) or paragraphs clearly separated by an empty space. However, while the song and the poem were printed in a poetic, line-by-line format, the prose format was presented as a running text. Of course, it needs to be acknowledged that not all songs or poems are presented this way. The present target song and poem, however, were arranged in a line-by-line format, leaving semantic units largely intact. While the present study cannot provide evidence regarding any possible benefits of poetic formatting, I want to

speculate on the possible impact of a line-by-line structure on the learners' processing of the text.

Firstly, it seems likely that a poetic format renders certain phonological features more salient and thus draws attention to linguistic form. Particularly, words in line-final and line-initial position might be more noticeable to the reader. As the chi-square analysis in the previous section showed, words that occurred in line-final position of the written song lyrics were better remembered than in the context of an uninterrupted running text in the prose format. Further research is necessary to determine whether this difference in recall was due to salient sound patterns such as end rhyme, to the formatting or both. Poetic devices such as alliteration might also become more apparent in shorter lines. As discussed in the previous section, greater salience of structural features of the text could then lead to processing and possibly increased retention of linguistic form.

Secondly, the visible presentation of words in written chunks might encourage the processing and encoding of language as larger units. Hartley (1993), for example, found that children reading in their native language recalled text as chunked if it was presented in a chunked, phrase-by-phrase format separated along semantic boundaries. However, no study has explored the effect of text segmentation on non-native readers.

And finally, the visible segmentation of text into semantic phrases might support reading as well as listening comprehension. Research has shown that poor reading comprehension is partially due to the readers' inability to perceive larger semantic units in the text (Jandreau & Bever, 1992). Frase and Schwartz (1979) reported experimental findings that L1-readers read more efficiently when lines represented "meaningful groups of information" (p. 205) and when semantic units were not interrupted by a line break. As we have seen above, attention to form is a prerequisite to learning (Schmidt, 1990, 2001), and L2-learners can only attend to form if the meaning of the input can be easily processed (VanPatten, 1989). However, a discussion of the effect of the written line-by-line format remains speculative. Further research is required to investigate the impact of the text format on verbatim retention.

6.5.3 Song-based teaching, cognitive elaboration and episodic memory

Another possible factor underlying the increased verbatim text retention following a song- and poem-based teaching unit is the multimodality of the input, which might result in more elaborated processing of the input and in a rich episodic memory to be formed. A song-based lesson includes various elements that encourage greater elaboration of the input, that is, in a greater number and variety of cognitive processes to occur (cf. Craik & Tulving, 1975). Songs comprise several stimuli that are processed at various levels of cognitive depth such as music, rhythm, linguistic sound patterns, mental imagery and also physical or kinaesthetic experiences associated with the perception and production of songs (Molnar-Szakacs & Overy, 2006). In addition, song lyrics often evoke a strong emotional response by means of the music (cf. Peretz & Zatorre, 2003) and also the emotionally evocative lyrics (cf. Murphey, 1989; 1990a). The lesson around a song or poem provides further contextual stimuli, including social factors such as the synchronised and affective experience of listening to and singing along with a song as a group (Overy & Molnar-Szakacs, 2009). The elaboration as well as the depth of processing can produce a strong episodic memory comprising various associated memory traces, which might also serve as associated retrieval cues during recall and recognition. As explained in Chapter 2, episodic memory refers to autobiographical memory of personally experienced events, which includes contextual information about content, time and place of the experience (Tulving, 1972). Episodic memory allows the "rememberer" to engage in mental time travel, that is, the mental revisiting of a past event (Tulving, 2002). Such mental time travel might assist the memorization and retrieval of words because the learner can re-experience the learning event including the verbal input encountered during that experience. In addition, Baddeley (1988) argued that semantic memory, including verbal knowledge, is information abstracted from learning events initially processed and stored as autobiographical memory. Strong episodic memory involving a variety of associated memory traces derived from a richly elaborated song-based teaching unit can form a solid basis for the formation of abstracted semantic memory. Finally, it can also draw the learners' attention to the target vocabulary outside the original learning context and thus increase the likelihood of a target item to be acquired, including the use of a word in different contexts and its use

in longer multi-word units. The latter point is explored in more detail in the next section on the pedagogical implications of these findings.

6.6 Pedagogical implications of the thesis

The findings from this thesis have implications for the use of songs in the language classroom. Firstly, the overall findings support teachers in their use of songs in class. The three studies together, that is, the survey study, the song corpus analysis and the quasiexperiment, provide empirical validation for the way language instructors tend to utilise songs as teaching materials in class, at least as far as the fostering of lexical learning is concerned. The use of songs as described by a majority of survey informants can result in increased verbatim text retention. As we have seen throughout this thesis, verbatim text retention can benefit lexical learning by means of facilitating the acquisition and consolidation of word knowledge and by assisting the acquisition of formulaic language, that is, multi-word units. As discussed in the previous section, findings from the intervention study imply that songs and also poems might draw the learners' attention to the surface structure of the text, thus encouraging noticing of formal aspects of the language and aiding the memorization of words and phrases. It appears, then, that the increased attention to and retention of surface form in both songs and poems can provide opportunities to facilitate the acquisition of deep word knowledge beyond a first form-meaning connection, including a word's grammatical functions, its collocations and its idiomatic uses. This use of songs is further supported by findings regarding the lexical profile of teacher-selected songs. As the song corpus study showed, songs, particularly when used with advanced learners, seem to offer limited learning opportunities in terms of new vocabulary. Instead, they provide opportunities to repeatedly encounter highfrequency words in the context of longer connected (and authentic) text and thus permit learners to notice the use of a word in context.

Results from the quasi-experimental intervention study further indicate that songs are an appropriate tool to foster the acquisition of multi-word units. On the recognition test, participants in the song (and also the poem) condition overall showed higher recognition of the original phrases from the target text, which predominantly differed from their paraphrases due to changes affecting form and not meaning. Such a memory effect

based on structural rather than semantic features might be particularly helpful for multiword items that cannot be explained on the basis of semantic reasoning such as etymological elaboration (Boers & Demecheleer, 1998; Boers, 2001). While some collocations can literally be explained, others offer little in terms of intrinsic semantic motivation. For example, teachers might find it difficult to give a reason why the English expression "to <u>take</u> the first step" translates into the German expression "*den ersten Schritt <u>machen</u>*" (literally, "to <u>make</u> the first step"). In such cases, a teaching and learning approach based on memorization and conscious awareness of form as exemplified by the song-based teaching unit implemented in the intervention study might make up for a lack of opportunities for semantic elucidation.

Overall, multi-word items are abundant in English and other languages, and both nativelike production as well as native-like reception and decoding of target language seem to depend to a large extent on the knowledge of formulaic language (Pawley & Syder, 1983). Consequently, second language learners need to learn a great number of multiword units in order to become both receptively and productively proficient in a second language. Given the abundance of fixed expressions, phrasal verbs and other multi-word units in a language, this is a challenging task that requires a great deal of verbatim memorization. Songs can offer an additional (and enjoyable) pathway to memorizing linguistic formulas and to learning and consolidating knowledge of multi-word units. The intervention study, which forms part of this thesis, has provided evidence that songs and poems can be used as a mnemonic aid to facilitate the memorization of word strings. Presenting formulaic sequences in a song or poem might thus assist in the enormous task of expanding language learners' "resource bank of multi-word expressions" (Boers & Lindstromberg, 2005, p. 234).

At this point, it should also be reiterated that a song-based lesson can result in a strong episodic memory comprising a variety of associated memory traces. If learners encounter a target word or expression outside the classroom, this might trigger retrieval of various connected memory traces of the song-based learning event, including the melody, emotions and contextual information of the classroom situation, and consequently render the word more noticeable during that new encounter. Such noticing of a word in a

new context can further support the consolidation of the target vocabulary, and it can foster the acquisition of further word knowledge due to the novel context.

While songs and poems seem to have intrinsic qualities that draw the listeners' and readers' attention towards the surface structure, resulting in higher retention of word form, I would like to propose that more explicit awareness-raising of structural elements might have resulted in even greater memorization of verbatim text. This assumption is supported by Boers, Lindstromberg and Eyckmans' (2014) finding that alerting learners to the presence of alliteration resulted in better retention of the alliterative phrases than their incidental learning during a meaning-focused activity. The authors emphasised that "awareness-raising has an important role to play if teachers and/or materials writers wish to help learners to take advantage of the memorability that alliteration appears to afford" (p. 7). It should, therefore, be considered that the use of songs (and poems) in class could be enhanced by explicitly pointing out sound patterns in the lyrics.

Another factor that might have further aided noticing of form under the song and the poem condition in the present study, was the written text and in particular its poetic format. The line-by-line presentation with its clear breakdown into short clauses and its visible segmentation into discernible semantic units might have increased the perceptual salience of phonemic patterns within the text and might thus have further aided the memorization of text. Consequently, it seems advisable to provide learners with a written copy of the (well laid-out) lyrics when using songs to teach formal aspects of the language.

Furthermore, testing learners' recall of the text might be appropriate not just in an experimental setting but also during regular language teaching and might be a feasible way of further motivating students to memorise elements of the target text. However, while participants in the present study were not alerted to the purpose of the memory tests following the treatment lesson, teachers should consider announcing the testing of verbatim memory, as this might increase learners' motivation to focus on form. Empirical research into verbatim memory of connected text indicates that retention rates are indeed higher when participants are warned about the fact that word-for-word memory will be tested (Gurevich et al., 2010).

Finally, existing research (e.g. Barcroft, 2002) indicates that a strong focus on linguistic form can come at the expense of processing semantic content. This implies that semantic processing and comprehension of lyrics during a song-based lesson might benefit from additional support. As findings from the survey study indicated, many teachers support their learners in the comprehension of a song in various ways. The frequent repetitions of a song during a lesson and the integration of meaning-focused activities into the song-based teaching unit provide opportunities for semantic elaboration and clarification. Overall, the survey results seem to indicate that songs lend themselves well to repeated listening and also singing. As a majority of informants observed, songs can be repeated several times without boring the students. In addition, respondents agreed that songs can be further re-encountered outside the classroom. The corpus study also attested to a high repetition of lexis in the song lyrics. The present thesis thus provides some support for the use of songs as enjoyable "drill" activities.

The aspect of enjoyment also needs to be further considered when deciding whether to use poems instead of songs. As the intervention study showed, the song-based and the poem-based treatment resulted in near-equivalent retention of verbatim text. However, survey respondents attested to the fact that songs are enjoyable, relaxing and interesting for the learners. Webb and Rodgers (2009b), when discussing the use of TV programmes for lexical learning, pointed out that "although we strongly advocate learning vocabulary through reading, the fact that people spend more time watching television than reading suggests that it could be an effective method of learning vocabulary" (p. 356). This claim also seems to hold true for pop songs, which are clearly more popular than poetry. Additionally, the "repeatability" of songs, the possibility to encounter them outside class, the emotional response evoked by music (Peretz & Zatorre, 2003; Krueger, 2011; Krueger, 2014) and also social factors such as the synchronised and affective experience of listening to and singing along with music as a group (Overy & Molnar-Szakacs, 2009) speak for the use of songs rather than poems.

Finally, the present thesis revealed an occasional conflict between teacher cognitions and actual classroom practices. The survey showed that many teachers view songs as a useful tool to foster language acquisition and want to use them in class. However, their intentions are often thwarted by a lack of official teaching material involving songs. As

the questionnaire highlighted, finding an appropriate song for a particular purpose, developing a lesson plan and designing complementary activities is perceived to take a lot of effort and is seen to be highly time-consuming. In addition, teachers occasionally encounter opposition from the school administration, from colleagues and from parents. The provision of official musical material would both validate teachers' choice to use songs for pedagogical purposes and facilitate the preparation and implementation of language lessons involving songs.

6.7 Original contributions of the study

The research presented in this thesis offers a number of original contributions to the field of language learning through songs. These include previously unexplored aspects of the research area as well as methodologies not previously employed to investigate the use and usefulness of songs in foreign language teaching. Firstly, the present study applied three very distinct methodological approaches which produced one coherent picture of song use and its effects on lexical learning. Secondly, the present thesis has explored songs as they are actually used. In order to do so, rich data was gathered on classroom practices as well as on the songs used as teaching material. The elicited information was integrated into the design of a quasi-experiment, thus ensuring a high level of ecological validity. Empirical research into lexical learning through songs is scarce and has mainly focused on complete or continuing beginners. The present intervention study adds further insight into the use of songs with intermediate and advanced learners. In addition, using a recognition test in the quasi-experimental intervention study provided an insight into small levels of memorization frequently neglected in research investigating lexical learning through songs. Finally, the quasi-experiment not only compared a sung and spoken learning condition, but also integrated a poem (as opposed to prose) condition, thus not only considering the melody as a factor underlying the mnemonic potential of songs but also exploring the effects of lexical sound patterns on the retention of words and longer connected text.

Furthermore, this project investigated the subject from the point of view of the teacher. Only three other studies have so far explored the views of teachers on the topic on a very small scale. Additionally, the current thesis is distinct from existing research in its detailed exploration of actual teaching practices involving songs. Other distinguishing

factors are the large number of questionnaire informants and their diverse backgrounds and settings. Informants were located in 41 countries and taught at a variety of institutions. Consequently, the questionnaire in this thesis was sensitive to potential cultural issues and particular institutional needs and constraints. The corpus study is a further original contribution to the field of inquiry. In order to explore the lexical load of teacher-selected songs as well as the vocabulary learning opportunities they afford, I applied the approach of lexical thresholds (Laufer, 1989) and vocabulary coverage to songs. While this approach has been used to assess the lexical load of a number of text genres, it had until now not been applied to songs. Its use to establish a lexical profile of teacher-selected songs, therefore, permits a comparison with other text genres used in the language classroom. Overall, this thesis contributes to the investigation of song use in the language classroom due to its detailed findings regarding teaching practices involving songs as well as the lexical profile of teacher-selected songs and the opportunities a song-based teaching approach affords for lexical learning.

6.8 Limitations of the thesis as a whole

The limitations of each study presented in this thesis are discussed in detail in the respective chapters. However, there are several limitations that apply to the research project as a whole. Most importantly, the present thesis combined three methodological approaches which can be either considered qualitative (questionnaire) or quantitative (corpus study, intervention study). In addition, this thesis combined two different frameworks with often diverging goals and priorities. The questionnaire on the one hand explored classroom practices and the underlying intrinsic and extrinsic factors that influence teachers' pedagogical choices. In other words, the questionnaire focused on "what language teachers think, know, believe and do" (Borg, 2003, p. 81) regarding the use of songs and can thus be located within the field of teacher cognition research. It produced rich data and a greater understanding of various aspects of song use for pedagogical purposes. The quasi-experimental intervention study on the other hand focused on the formation of episodic memory and retention of verbatim text and can be located within a framework of cognition and memory research. Cognitive-psychological memory research is typically defined by its strictly controlled conditions and clear limitation of variables. Consequently, while incorporating a large number of findings from

the teacher questionnaire and also the corpus study, the quasi-experiment had the single purpose to ascertain whether a song used within a certain naturalistic setting resulted in higher recall and recognition of text compared to other formats. The present research study, thus, had to strike a balance between ecological validity and experimental control. As a consequence, the quasi-experiment focused on only one, that is, the most typical implementation of songs in the EFL classroom. However, it needs to be acknowledged that song-based teaching can take a variety of forms, as the diverse survey results showed, and that the findings of this study might not apply to other ways of using songs in class.

It also needs to be highlighted that the ecological validity of the intervention implemented during the quasi-experimental study reflected typical teaching practices involved in song-based teaching, as investigated through a teacher survey. However, the way the comparison treatments were designed was not informed by such a survey, and so it cannot be guaranteed that the ways the students in the poem and prose conditions dealt with the text were also "typical". Still, care was taken that no teaching practices were used that can be considered inappropriate for these other two text formats (see Chapter 5 for a detailed description of the treatment tasks).

Another caveat of the present study was the limited control of individual learner differences and their potential interaction with the learning conditions and various other aspects of the intervention. More importantly still, the text formats comprised a number of potential variables such as rhythm, metre and phonemic features. These were considered in combination only, defined overall as "text format". As a consequence, the intervention study was able to show a mnemonic effect of the text formats song and poem overall, but it was unable to discern, at least at the present level of analysis, which characteristics were pivotal in producing such a memory effect. However, further research and a more in-depth analysis of the test data could possibly shed further light on various moderating variables connected to the melody, the rhythm and textual features.

Another limitation that needs to be mentioned in this context is the use of only one song for the intervention study. It would have been more advisable to use several songs. In

addition, a counter-balanced design would have allowed for a greater control of individual learner differences, as it would have permitted for all participants to engage in all learning conditions at some point.

Furthermore, in this thesis learning was defined as memorization of verbatim text and measured by means of recall and recognition tests. While memorizing the form of lexical items is a prerequisite to using them, this thesis did not explore to what extent participants were able to receptively or productively use the memorised language. Recall and recognition of verbatim text is part of episodic memory, and it remains unclear whether participants also formed semantic memory of targeted lexical items, which would enable them to use the language dissociated from the original learning context. Finally the study was limited to the perspective of the teacher and did not explore the views of language learners.

6.9 Future research

While the use of songs in second language teaching is discussed at length in the pedagogical literature, only few empirical studies have investigated song's use and usefulness in language learning. The present study has provided some insight into current teaching practice involving songs. Furthermore, it has contributed to the understanding of the lexical profile and vocabulary load of teacher-selected songs. Finally, it has provided some evidence that songs can benefit lexical learning, particularly the acquisition and consolidation of linguistic form, due to a mnemonic effect of songs.

However, many questions remain unanswered. In particular, the mnemonic effect of songs as well as poems particularly regarding formal aspects of the target text should be further explored. More controlled experimental research is needed to explore what factors contribute to a greater memorability of songs and poems. Firstly, the present study has indicated that rhythm, metre and phonemic features such as alliteration, assonance, consonance and end rhyme might have contributed to a higher retention of the song lyrics and the poem, because these text modalities privilege learners' attention to such features. The effect of these variables on verbatim retention separately as well as in combination need to be further explored under more controlled conditions. Furthermore, future research should consider text-internal variables such as the number

of repetitions of words and phrases and their location within the text. Hyman and Rubin (1990), for example, found a primacy effect as well as a significant effect of line repetition regarding verbatim retention of Beatles' lyrics by native speakers. In addition, factors pertaining to the memorised language itself and their impact on retention need to be investigated, including such characteristics as the semantic concreteness of words, the semantic transparency of multi-word units, the frequency of words and phrases in the target language overall, and a possible equivalence of formulaic sequences in the native language of learners.

Another area that has received little attention in existing research is learners' ability to use the memorised language outside the original context. Testing verbatim recall and recognition of target text measures the formation of episodic memory, but falls short of measuring the learners' ability to productively use the memorised language in a new context. Ludke (2010) studied the retention of phrases in an unknown language in a sung and a spoken condition. Her test battery included a production test and a so-called conversation test. However, these tests required productive use of target items only within the original semantic context and in response to stimuli encountered in a pairassociate format. Given the lack of research, further investigation into the actual use of language learnt through songs is warranted.

Another relevant area is the comprehension of text when presented in the format of a song. As the questionnaire has revealed, form-focused teaching and in particular the teaching of vocabulary is only one of many uses of songs in the language classroom. In fact, teachers frequently use songs to provide meaningful input and to encourage learners to produce meaningful output in spoken and written form. At the same time, this study has provided some evidence that songs draw attention to the surface structure of the language. Consequently, further research into learners' comprehension of song lyrics is warranted. In this context, it needs to be highlighted that ideal vocabulary coverage of song lyrics for adequate comprehension has still to be established.

Finally, the present study has provided some tentative indication that the written format of a text might affect verbatim text retention, particularly in conjunction with a highly salient surface structure and when the target text is both listened to and read at the

same time. The segmentation of written text into smaller semantic units that also highlight structural features might thus be an additional way to aid language learners' comprehension and/or memorization of text and vocabulary. The written format has, unfortunately, received very little attention in second language research overall and in connection with the use of songs in language teaching in particular. However, it could prove to be an effective way to improve written teaching materials, and should thus be further investigated.

6.10 Conclusion

The current thesis explored the use of songs in the language classroom and their usefulness for lexical learning by means of investigating three aspects relevant to songbased teaching: the teacher, the material and the learning outcome. It combined an empirical research approach and a clear focus on classroom reality and integrated both qualitative and quantitative research methods. More specifically, it investigated teacher cognitions and teaching practices on the one hand, the lexical profile of songs used as teaching materials on the other hand and finally also the effects of song-based teaching on the retention of verbatim text. The findings of the present research project have provided evidence that the way teachers currently tend to use songs in language classrooms around the world can benefit lexical learning, in particular the consolidation of already familiar vocabulary and the acquisition of deep word knowledge beyond a first form-meaning connection. This project was also able to show that songs selected by teachers and material designers for in-class use overall meet the lexical needs of L2learners and can serve as an "entry-level" authentic text genre, particularly in combination with current teaching practices typical for a song-based lesson. At this point in time, songs are still often seen as teaching material located outside the official box of pedagogical tools. The present research findings, however, indicate that songs can facilitate vocabulary acquisition. This thesis has, thus, provided empirical validation for the way many teachers tend to use songs in language teaching and has given further evidence for the potential of songs for language learning, which has yet to be fully exploited.

References

- Abbott, M. (2002). Using music to promote L2 learning among adult learners. *TESOL Journal*, *11*(1), 10-17.
- Abrate, J. H. (1983). Pedagogical applications of the French popular song in the foreign language classroom. *The Modern Language Journal, 67*(1), 8-12.
- Adkins, S. (1997). Connecting the powers of music to the learning of languages. Document Resume, 45.
- Adolphs, S., & Schmitt, N. (2003). Lexical coverage of spoken discourse. *Applied Linguistics*, *24*(4), 425-438+557.
- Anton, R. J. (1990). Combining singing and psychology. *Hispania*, 1166-1170.
- Aquil, R. (2012). Revisiting songs in language pedagogy. *Journal of the National Council of Less Commonly Taught Languages, 11,* 75-95.
- Arleo, A. (2000). Music, song and foreign language teaching. *Les Cahiers de l'APLIUT,* 19(4), 5-19.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *Psychology of learning and motivation* (Vol. 2, pp. 89-195). New York: Academic Press.
- Baddeley, Alan D. (1988). Cognitive psychology and human memory. *Trends in Neurosciences, 11*(4), 176-181.
- Baddeley, Alan D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417-423.
- Baddeley, A. D., Eysenck, M. W., & Anderson, M. (2009). *Memory*. New York; Hove: Psychology Press.
- Baddeley, A. D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*(1), 158-173.
- Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language*, *61*(3), 438-456.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation* (pp. 47-89). New York: Academic Press.
- Baechtold, S., & Algier, A. (1986). Teaching college students vocabulary with rhyme, rhythm, and ritzy characters. *Journal of Reading*, *30*(3), 248-253.
- Barber, E. J. W. (1980). Language acquisition and applied linguistics. *ADFL Bulletin*, 12(1), 26-32.
- Barcroft, J. (2002). Semantic and structural elaboration in L2 lexical acquisition. *Language Learning*, *52*(2), 323-363.
- Barcroft, J. (2004). Second language vocabulary acquisition: A lexical input processing approach. *Foreign Language Annals, 37*(2), 200-208.
- Barcroft, J. (2007). When knowing grammar depends on knowing vocabulary: Nativespeaker grammaticality judgements of sentences with real and unreal words. *The Canadian Modern Language Review / La revue canadienne des langues vivantes,* 63(3), 313-343.
- Bartlett, F. C. (1932). *Remembering: An experimental and social study*. London: Cambridge University Press.
- Bartlett, J. C., & Snelus, P. (1980). Lifespan memory for popular songs. *The American Journal of Psychology*, *93*(3), 551-560.
- Bauer, L., & Nation, P. (1993). Word families. *International Journal of Lexicography*, 6(4), 253-279.
- Bieber, J., Stewart, C., Nash, T., Bridges, C., & Milian, C. (2010). Baby (performed by Justin Bieber). My World 2.0 (CD single/digital download). London: Island.
- Bjorklund, K. S. (2002). *Music application in the ESL classroom*. Unpublished doctoral thesis, University of Wyoming, Laramie, USA.
- Boers, F. (2001). Remembering figurative idioms by hypothesizing about their origin. *Prospect, 16*(3), 34-43.
- Boers, F., & Demecheleer, M. (1998). A cognitive semantic approach to teaching prepositions. *ELT Journal*, *52*(3), 197-204.

- Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research*, *10*(3), 245-261.
- Boers, F., & Lindstromberg, S. (2005). Finding ways to make phrase-learning feasible: The mnemonic effect of alliteration. *System*, *33*(2), 225-238.
- Boers, F., & Lindstromberg, S. (2009). *Optimizing a lexical approach to instructed second language acquisition*. New York: Palgrave Macmillan.
- Boers, F., Lindstromberg, S., & Webb, S. (2014). Further evidence of the comparative memorability of alliterative expressions in second language learning. *RELC Journal*, *45*(1), 85-99.
- Boers, F., Lindstromberg, S. & Eyckmans, J. (2014). Is alliteration mnemonic without awareness-raising? *Language Awareness*, *23*(4), 291-303.
- Bonk, W. (2000). Second language lexical knowledge and listening comprehension. International Journal of Listening, 14(1), 14-31.
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, *36*(02), 81-109.
- Bradburn, N. M., Wansink, B. & Sudman, S. (2004). Asking questions: The definite guide to questionnaire design for market research, political polls, and social and health questionnaires. San Francisco, CA: Jossey-Bass.
- Brown, R., Waring, R. & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, *20*(2), 136-163.
- Calvert, S. L., & Billingsley, R. L. (1998). Young children's recitation and comprehension of information presented by songs. *Journal of Applied Developmental Psychology*, 19(1), 97-108.
- Calvert, S. L., & Tart, M. (1993). Song versus verbal forms for very-long-term, long-term, and short-term verbatim recall. *Journal of Applied Developmental Psychology, 14*(2), 245-260.

- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*(3), 354.
- Chambers, C. K. (1980). *Non-lexical vocables in Scottish traditional music.* Unpublished doctoral thesis, University of Edinburgh, Edinburgh, UK.
- Chapple, L., & Curtis, A. (2000). Content-based instruction in Hong Kong: Student responses to film. *System*, *28*(3), 419-433.
- Chazin, S., & Neuschatz, J. S. (1990). Using a mnemonic to aid in the recall of unfamiliar information. *Perceptual and Motor Skills*, *71*(3), 1067-1071.
- Claerr, T. A., & Gargan, R. (1984). The role of songs in the foreign language classroom. *OMLTA Journal, 28*, 32.
- Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, *29*(1), 72-89.
- Conrad, B. (1991). Contemporary and historical contexts in popular music: Theory and implementation of an intermediate level course. *Die Unterrichtspraxis/Teaching German*, 18-30.
- Coxhead, A. (2000). A new academic word list. TESOL Quarterly, 34(2), 213-238.
- Craik, F. I. (2002). Levels of processing: Past, present... and future? *Memory*, *10*(5), 305-318.
- Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671-684.
- Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268-294.
- Dethier, B. (1991). Using music as a second language. English Journal, 72-76.
- Domoney, L., & Harris, S. (1993). Justified and ancient: Pop music in EFL classrooms. *ELT Journal*, *47*(3), 234-241.

- Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in second language research: Construction, administration, and processing*. New York; London: Routledge.
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.
- Dowling, W. J., Tillman, B., & Ayers, D. F. (2001). Memory and the experience of hearing music. *Music Perception: An Interdisciplinary Journal, 19*(2), 249-276.
- Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research, 26*(2), 163-188.
- Ebbinghaus, H. (1885). Über das Gedächtnis: Untersuchungen zur experimentellen Psychologie. Leipzig: Duncker & Humblot.
- Edwards, J. C. (1997). Using music for second language purposes. Unpublished Master's thesis, California State University, Dominguez Hills, USA.
- Ellis, E. M. (2006). Language learning experience as a contributor to ESOL teacher cognition. *TESL-EJ: Teaching English as a Second or Foreign Language, 10*(1), 1.
- Engh, D. (2013). Why use music in English language learning? A survey of the literature. English Language Teaching, 6(2), p113.
- Esa, M. (2008). Musik im Deutschunterricht: Der gezielte Einsatz. *Die Unterrichtspraxis/Teaching German, 41*(1), 1-14.
- Eschrich, S., Münte, T. F., & Altenmüller, E. O. (2008). Unforgettable film music: The role of emotion in episodic long-term memory for music. *BMC Neuroscience*, *9*(1), 48.
- Ferreri, L., Aucouturier, J., Muthalib, M., Bigand, E., & Bugaiska, A. (2013). Music improves verbal memory encoding while decreasing prefrontal cortex activity: An fNIRS study. *Frontiers in Human Neuroscience*, 7, 1-9.
- Flesch, R. (1974). The art of readable writing. New York: Harper & Row.
- Foncesa Mora, C. (2000). Foreign language acquisition and melody singing. *ELT journal,* 54(2), 146-152.
- Frase, L. T., & Schwartz, B. J. (1979). Typographical cues that facilitate comprehension. Journal of Educational Psychology, 71(2), 197.

- Gardner, R. C., Tremblay, P. F., & Masgoret, A. (1997). Towards a full model of second language learning: An empirical investigation. *The Modern Language Journal*, *81*(3), 344-362.
- Gatti-Taylor, M. (1980). Songs as a linguistic and cultural resource in the intermediate Italian class. *Foreign Language Annals, 13*(6), 465.
- Gelman, M. (1973). Poetry and songs in the teaching of languages. *Babel, 9*(1), 13-15.
- Ginsborg, J., & Sloboda, J. A. (2007). Singers' recall for the words and melody of a new, unaccompanied song. *Psychology of Music, 35*(3), 421-440.
- Green, J. M. (1993). Student attitudes toward communicative and non-communicative activities: Do enjoyment and effectiveness go together? *The Modern Language Journal*, 77(1), 1-10.
- Greenberg, D. L., Keane, M. M., Ryan, L., & Verfaellie, M. (2009). Impaired category fluency in medial temporal lobe amnesia: The role of episodic memory. *Journal of Neuroscience*, 29(35), 10900-10908.
- Greenberg, D. L., & Verfaellie, M. (2010). Interdependence of episodic and semantic memory: Evidence from neuropsychology. *Journal of the International Neuropsychological Society*, 16(5), 748-753.
- Griffee, D. (1988). Song and music rechniques in foreign and second language classrooms. *Cross Currents, A Journal of Language Teaching and Cross-Cultural Communication, 15*(1), 12.
- Gruber, K.J., & Gaebelein, J. (1979). Sex differences in listening comprehension. *Sex Roles,* 5(3), 299-310.
- Gurevich, O., Johnson, M. A., & Goldberg, A. E. (2010). Incidental verbatim memory for language. *Language and Cognition*, 2(1), 45-78.
- Hahn, S. M. (1972). The effect of music in the learning and retention of lexical items in German. Retrieved from http://eric.ed.gov/?id=ED119455.
- Halpern, A. R. (2001). Cerebral substrates of musical imagery. *Annals of the New York Academy of Sciences, 930*(1), 179-192.

- Halpern, A. R., & Bartlett, J. C. (2011). The persistence of musical memories: A descriptive study of earworms. *Music Perception*, 28(4), 425-431.
- Hartley, J. (1993). Recalling structured text: Does what goes in determine what comes out? *British Journal of Educational Technology*, *24*(2), 84-91.
- Heatley, A., Nation, P., & Coxhead, A. (2002). RANGE and FREQUENCY programs. Available at http://www.victoria.ac.nz/lals/staff/paul-nation.aspx.
- Heyworth, F. (2006). The common European framework. ELT Journal, 60(2), 181-183.
- Hildred, K. (2011). Grammar and rhythm: Making the most of music in the English language classroom. *EA Journal*, *27*(1), 55-58.
- Hilton, H. (2008). The link between vocabulary knowledge and spoken L2 fluency. *The Language Learning Journal, 36*(2), 153-166.
- Hirsh, D, & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, *8*(2), 689-696.
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125-132.
- Hu, M., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language, 13*(1), 403-430.
- Huy Lê, M. (1999). The role of music in second language learning: A Vietnamese perspective. A paper presented at the combined 1999 Conference of the Australian Association for Research in Education and the New Zealand Association for Research.
- Hwang, K. , & Nation, P. (1989). Reducing the vocabulary load and encouraging vocabulary learning through reading newspapers. *Reading in a Foreign Language*, 6(1), 323.
- Hyman, I. E., & Rubin, D. C. (1990). Memorabeatlia: A naturalistic study of long-term memory. *Memory & Cognition, 18*(2), 205-214.
- Jandreau, S., & Bever, T. G. (1992). Phrase-spaced formats improve comprehension in average readers. *Journal of Applied Psychology*, 77(2), 143.

- Jarvella, R. J. (1971). Syntactic processing of connected speech. *Journal of Verbal Learning* and Verbal Behavior, 10(4), 409-416.
- Jey, J., Lobina, M., & Gabutti, M. (1998). Blue (da ba dee) (performed by Eiffel 65). *Europop (CD single/digital download)*. New York: Universal Records.
- Johnson, K. E. (1994). The emerging beliefs and instructional practices of preservice English as a second language teachers. *Teaching and Teacher Education*, 10(4), 439-452.
- Jolly, Y. S. (1975). The use of songs in teaching foreign languages. *The Modern Language Journal*, *59*(1-2), 11-14.
- Kanel, K. (1997). Teaching with music: A comparison of conventional listening exercises with pop song gap-fill exercises. *JALT Journal*, *19*(2), 217-234.
- Kao, T., & Oxford, R. L. (2014). Learning language through music: A strategy for building inspiration and motivation. System, 43, 114-120.
- 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.
- Kilgour, A. R., Jakobson, L. S., & Cuddy, L. L. (2000). Music training and rate of presentation as mediators of text and song recall. *Memory & Cognition, 28*(5), 700-710.
- Klahr, D., Chase, W. G., & Lovelace, E. A. (1983). Structure and process in alphabetic retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 9(3), 462-477.
- Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal, 73*(4), 440-464.
- Krashen, S. (1983). The Din in the head, input, and the language acquisition device. Foreign Language Annals, 16(1), 41-44.
- Krashen, S., & Terrell, T. D. (1983). *The natural approach: Language acquisition in the classroom*. San Francisco: Pergamon Press.

- Kreyer, R., & Mukherjee, J. (2007). The style of pop song lyrics: A corpus-linguistic pilot study. *Anglia Zeitschrift für englische Philologie, 125*(1), 31-58.
- Kroeger, C. (2008). If today was your last day (recorded by Nickelback). *Dark Horse* (*digital download*). New York: Roadrunner.
- Krueger, J. (2011). Doing things with music. *Phenomenology and the Cognitive Sciences, 10*(1), 1-22.
- Krueger, J. (2014). Affordances and the musically extended mind. *Frontiers of Psychology*, 4, 1003-1015.
- Lake, R. (2002). Enhancing acquisition through music. *The Journal of the Imagination in Language Learning and Teaching*, *7*, 98-106.
- Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), Special language: From humans thinking to thinking machines (pp. 316-323). Clevedon: Multilingual Matters.
- Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics, 19*(2), 255-271.
- Laufer, B. (2001). Quantitative evaluation of vocabulary: How it can be done and what it is good for. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, T.
 McNamara and KO'Loughlin (*Eds.*), *Experimenting with uncertainty* (pp. 241-250).
 Cambridge: Cambridge University Press.
- Laufer, B. (2013). Lexical thresholds for reading comprehension: What they are and how they can be used for teaching purposes. *TESOL Quarterly*, *47*(4), 867-872.
- Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning*, *54*(3), 399-436.
- Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15-30.
- Laufer, B., & Shmueli, K. (1997). Memorizing new words: Does teaching have anything to do with it? *RELC Journal, 28*(1), 89-108.

- Lems, K. (2005). Music works: Music for adult English language learners. *New Directions* for Adult and Continuing Education, 13-21.
- Levitin, D. J. (2006). *This is your brain on music: The science of a human obsession*. New York: Dutton.
- Lewis, M. (1993). The Lexical Approach: The state of ELT and a way forward. Hove: LTP.
- Lewis, M. (2002). *Implementing the Lexical Approach: Putting theory into practice*. Boston, MA: Thomson Heinle.
- Liikkanen, L. A. (2012). Musical activities predispose to involuntary musical imagery. *Psychology of Music, 40*(2), 236-256.
- Lindstromberg, S., Boers, F., & Eyckmans, J. (2014). Is alliteration mnemonic without awareness-raising? *Language Awareness*, *23*(4), 291-303.
- Lorenzutti, N. (2014). Beyond the gap fill: Dynamic activities for song in the EFL classroom. *English Teaching Forum, 1*, 14-21.
- Ludke, K. M. (2010). *Songs and singing in foreign language learning*. Unpublished doctoral thesis, University of Edinburgh, Edinburgh, UK.
- Lüke, M. (2008). Modern classics: Reflections on Rammstein in the German class. *Die Unterrichtspraxis/Teaching German, 41*(1), 15-23.
- Lund, R. J. (1991). A comparison of second language listening and reading comprehension. *The Modern Language Journal, 75*(2), 196-204.
- Markham, P. L. (1988). Gender and the perceived expertness of the speaker as factors in ESL listening recall. *TESOL Quarterly*, *22*(3), 397-406.
- McCauley, J. K., & McCauley, D. S. (1992). Using choral reading to promote language learning for ESL students. *The Reading Teacher*, *45*(7), 526-533.
- McElhinney, M., & Annett, J. M. (1996). Pattern of efficacy of a musical mnemonic on recall of familiar words over several presentations. *Perceptual and Motor Skills*, 82(2), 395.
- Meara, P. (1995). The importance of an early emphasis on L2 vocabulary. *The Language Teacher, 19*(2), 8-11.

- Medina, S. L. (1990). The effects of music upon second language vocabulary acquisition. Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages (San Francisco, CA, March 1990).
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81.
- Milton, J. (2008). Vocabulary uptake from informal learning tasks. *Language Learning Journal*, *36*(2), 227-237.
- Milton, J., & Meara, P. (1998). Are the British really bad at learning foreign languages? Language Learning Journal, 18(1), 68-76.
- Mizzy, V. (2010). Addams Family theme. *The Addams Family Original Music from the T.V. Show (digital download)*. New York: Legacy Recordings.
- Molnar-Szakacs, I., & Overy, K. (2006). Music and mirror neurons: From motion to 'e'motion. *Social cognitive and affective neuroscience*, 1(3), 235-241.
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*(5), 519-533.
- Moussard, A., Bigand, E., Belleville, S., & Peretz, I. (2012). Music as an aid to learn new verbal information in Alzheimer's Disease. *Music Perception: An Interdisciplinary Journal, 29*(5), 521-531.
- Murphey, T. (1987). English through music: A sheltered subject matter language course. Paper presented at the Bulletin CILA (Commission interuniversitaire suisse de linguistique appliquée).
- Murphey, T. (1989). The when, where, and who of pop lyrics: The listener's prerogative. *Popular Music, 8*(02), 185-193.
- Murphey, T. (1990a). Song and music in language learning. *European University Studies: Education. Bern: Peter Lang.*
- Murphey, T. (1990b). The song stuck in my head phenomenon: A melodic Din in the LAD? *System, 18*(1), 53-64.

Murphey, T. (1992). The discourse of pop songs. *TESOL Quarterly*, 26(4), 770-774.

- Murphey, T., & Alber, J. (1985). A pop song register: The motherese of adolescents as affective foreigner talk. *TESOL Quarterly, 19*(4), 793-795.
- Nation, P. (2004). A study of the most frequent word families in the British National
 Corpus. In P. Bogaards & B. Laufer-Dvorkin (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 3-13). Amsterdam, Philadelphia:
 John Benjamin Publishers.
- Nation, I.S.P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review/La revue canadienne des langues vivantes, 63*(1), 59-82.
- Nation, I.S.P. (2013). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, P. (2014). How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language, 26*(2), 1-16.
- Nelson, T. O. (1977). Repetition and depth of processing. *Journal of Verbal Learning and Verbal Behavior, 16*(2), 151-171.
- Newham, P. (1995). Making a song and dance: The musical voice of language. *Journal of the Imagination in Language Learning*, *3*, 66-74.
- Nini, A., & Dor, G. (2000). Blue touches blue (performed by Noa). *Blue touches blue (CD album)*. London: Polydor.
- North, A. C., Hargreaves, D. J., & O'Neill, S. A. (2000). The importance of music to adolescents. *British Journal of Educational Psychology*, *70*(2), 255-272.
- Overy, K., & Molnar-Szakacs, I. (2009). Being together in time: Musical experience and the mirror neuron system. *Music Perception: An Interdisciplinary Journal, 26*(5), 489-504.
- Page, J., & Plant, R. (1971). Stairway to heaven (performed by Led Zeppelin). *Led Zeppelin IV (CD album)*. New York: Atlantic Records.

- Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford: Oxford University Press.
- Palmer, C., Jungers, M. K., & Jusczyk, P. W. (2001). Episodic memory for musical prosody. *Journal of Memory and Language*, 45(4), 526-545.
- Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J.C. Richards & R.W. Schmidt (Eds.), *Language and communication* (pp. 191-225). New York: Longman.
- Peregoy, S. F., Boyle, O., & Cadiero-Kaplan, K. (2008). *Reading, writing and learning in ESL: A resource book for teaching K-12 English learners*. Boston: Pearson.
- Peretz, I., & Zatorre, R. J. (2003). *The cognitive neuroscience of music.* Oxford: Oxford University Press.
- Pérez Aldeguer, S., & Leganés Lavall, E. N. (2012). La música como herramienta interdisciplinar: Un análisis cuantitativo en el aula de lengua extranjera de primaria. *Revista de Investigación en Educación, 10*(1), 127-143.
- Postman, L. (1963). Does interference theory predict too much forgetting? *Journal of Verbal Learning and Verbal Behavior, 2*(1), 40-48.
- Purcell, J. M. (1992). Using songs to enrich the secondary class. *Hispania*, 75(1), 192-196.
- Purnell-Webb, P., & Speelman, C. P. (2008). Effects of music on memory for text. *Perceptual and Motor Skills, 106*(3), 927-957.
- Qualtrics. (2011). Qualtrics. Provo, Utah, USA: Qualtrics.
- Racette, A., & Peretz, I. (2007). Learning lyrics: To sing or not to sing? *Memory & Cognition*, 35(2), 242-253.
- Rainey, D. W., & Larsen, J. D. (2002). The effect of familiar melodies on initial learning and long-term memory for unconnected text. *Music Perception: An Interdisciplinary Journal, 20*(2), 173-186.
- Reeve, C, & Williamson, J. (1987). Look what you've done to my song. *Modern English Teacher, 14*(4), 33-36.

Repovš, G., & Baddeley, A. D. (2006). The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience*, *139*(1), 5-21.

Richards, J. (1969). Songs in language learning. Tesol Quarterly, 3(2), 161-174.

- Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology*, *39*(1), 475-543
- Riddiford, N. (1998). *Song talk*. Wellington: National Association of ESOL Home Tutor Schemes.
- Rubin, D. C. (1995). *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes*. New York: Oxford University Press.
- Rubin, D. C., & Wallace, W. T. (1989). Rhyme and reason: Analyses of dual retrieval cues. *Journal of Experimental Psychology*, 15(4), 698.
- Rubin, D. C., Wallace, W. T., & Houston, B. C. (1993). The beginnings of expertise for ballads. *Cognitive Science*, *17*(3), 435-462.
- Rubin, J. (1994). A review of second language listening comprehension research. *The Modern Language Journal, 78*(2), 199-221.
- Sacks, O. (2010). *Musicophilia: Tales of music and the brain.* New York: Random House.
- Salas, R. (2006). *Wordsetting.* An unpublished Master's thesis, Victoria University of Wellington, Wellington, New Zealand.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 129-158.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.
- Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. Language Teaching Research, 12(3), 329-363.
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal, 95*(1), 26-43.
- Schmitt, N., & Boyd Zimmerman, C. (2002). Derivative word forms: What do learners know? *TESOL Quarterly*, *36*(2), 145-171.

Scott, M. (2008). WordSmith Tools version 5. Liverpool: Lexical Analysis Software.

- Serafine, M. L., Crowder, R. G., & Repp, B. H. (1984). Integration of melody and text in memory for songs. *Cognition*, *16*(3), 285-303.
- Serafine, M. L., Davidson, J., Crowder, R. G., & Repp, B. H. (1986). On the nature of melody-text integration in memory for songs. *Journal of Memory and Language*, 25(2), 123-135.
- Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *The Quarterly Journal of Experimental Psychology*, 22(2), 261-273.
- Shiffrin, R. M. (1999). 30 years of memory. In C. Izawa (Ed.), On human memory: Evolution, progress, and reflections on the 30th anniversary of the Atkinson-Shiffrin model (pp. 17-33). Mahwah, NJ: Lawrence Erlbaum Associates.
- Smith Salcedo, C. (2002). *The effects of songs in the foreign language classroom on text recall and involuntary mental rehearsal.* Unpublished doctoral thesis, Louisiana State University, Baton Rouge, USA.
- Speelman, C. P., Sibma, S., & MacLachlan, S. (2012). Memory for sequence order in songs.
 Paper presented at the 12th International Conference on Music Perception and
 Cognition and the 8th Triennial Conference of the European Soceity for the
 Cognitive Sciences of Music, Thessaloniki, Greece.
- Sprenger, M. (1999). Learning & memory: The brain in action. Alexandria: ASCD.
- Stæhr, L. (2008). Vocabulary size and the skills of listening, reading and writing. *The Language Learning Journal*, *36*(2), 139-152.
- Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition, 31*(04), 577-607.
- Stansell, J. W. (2005). The use of music for learning languages: A review of the literature. [WWW document.] Retrieved 22/11/2014, from http://mste.illinois.edu/courses/ci407su02/students/stansell/Literature%20Revie w%201.htm.

- Stocker, C. (1923). Teaching French through folk songs. *The Modern Language Journal,* 7(5), 285-288.
- Thorndike, E. (1920). A consistent error in psychological ratings. *Journal of Applied Psychology, 4*(1), 25-29.
- Thorndike, E., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York: Teachers College, Columbia University.
- Tillmann, B., & Dowling, W. J. (2007). Memory decreases for prose, but not for poetry. *Memory & Cognition, 35*(4), 628-639.
- Timberlake, J., Mosley, T., & Hills, N. (2006). What goes around ... comes around (performed by Justin Timberlake). *FutureSex/LoveSounds (CD single/digital download)*. New York: Jive Records.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), Organization of memory (pp. 381-403). New York: Academic Press.
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie canadienne*, *26*(1), 1-12.
- Tulving, E. (1995). Organization of memory: Quo vadis? In M. S. Gazzaniga & E. Bizzi (Eds.), *The cognitive neurosciences*. Cambridge, MA: MIT Press.
- Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology,* 53, 1-25.
- Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5*(4), 381-391.
- van Asselt, J. (1971). Learning German through rhyme, rhythm and melody. *Die Unterrichtspraxis / Teaching German, 4*(1), 131-133.
- van Zeeland, H., & Schmitt, N. (2013). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics*, 34(4), 457-479.
- VanPatten, B. (1989). Can learners attend to form and content while processing input? *Hispania*, 72(2), 409-417.

- Walklett, C. (2014). *Changing attitudes to using songs and song lyrics in EFL*. Paper presented at the 48th Annual International IATEFL Conference and Exhibition, Harrogate, UK.
- Wallace, W. T., & Rubin, D. C. (1991). Characteristics and constraints in ballads and their effects on memory. *Discourse Processes*, *14*(2), 181-202.
- Wallace, W. T. (1994). Memory for music: Effect of melody on recall of text. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1471-1485.
- Webb, S., & Rodgers, M. P. H. (2009a). The exical coverage of movies. *Applied Linguistics*, 407-427.
- Webb, S., & Rodgers, M. P. H. (2009b). Vocabulary demands of television programs. *Language Learning*, *59*(2), 335-366.
- Welch, C. (2007). Choral music in the junior high/middle school: Teaching choral literature with foreign language texts. *Choral Journal*, *47*(10), 59-60.
- West, M. (1953). A general service list of English words: with semantic frequencies and a supplementary word-list for the writing of popular science and technology.
 Harlow, Essex: Longman.
- Williamson, V. J., & Jilka, S. R. (2014). Experiencing earworms: An interview study of Involuntary Musical Imagery. *Psychology of Music,* 42(5), 653-670.
- Williamson, V. J., Jilka, S. R., Fry, J., Finkel, S., Müllensiefen, D., & Stewart, L. (2012). How do "earworms" start? Classifying the everyday circumstances of Involuntary Musical Imagery. *Psychology of Music, 40*(3), 259-284.
- Winehouse, A. (2006). Rehab (performed by Amy Winehouse). *Back to Black [digital download]*. London: Island.
- Woods, D. (1996). *Teacher cognition in language teaching: Beliefs, decision-making, and classroom practice*. New York; Cambridge: Cambridge University Press.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge, MA: Cambridge University Press.

- Yalch, R. F. (1991). Memory in a jingle jungle: Music as a mnemonic device in communicating advertising slogans. *Journal of Applied Psychology, 76*(2), 268.
- Zatorre, R. J., & Salimpoor, V. N. (2013). Why music makes our brain sing: Op-Ed, *New York Times,* p. SR.12.
- Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology*. New York: Hafner.

Appendix

Appendix A: Materials for Study 1 (the teacher survey)

- Appendix A.1: Complete questionnaire (paper-based version)
- Appendix A.2: Ethics approval for the teacher survey

Appendix A.1: The teacher questionnaire

[Certain elements of the online survey such as drop-down menus, scroll-over texts and text boxes are not displayed in this paper version. This includes additional information for the respondents such as a definition of the class types "second language course" and "foreign language course", information about the proficiency levels, the recurrent display of the researcher's contact details and a progress bar indicating in percent how much of the survey participants have already completed and how much is yet to come. Also, scales, both vertical and horizontal, are displayed in a different layout.]

The Use (or Non-Use) of Songs in the Language Classroom – a Teacher Survey

My name is Friederike Tegge, and I am a PhD student at Victoria University of Wellington, New Zealand. I am carrying out a research project on the use of songs in the language classroom and I would like you to tell me your opinion on and approach to using songs. For that purpose I have put together the following questionnaire. It will help gain an insight into how and **why** – or **why not** – language teachers use songs in the language classroom.

This survey is anonymous. No opinions will be attributed to you in any way that will identify you. Access to the research data is restricted to me and my two supervisors. The data may contribute to research publications. It will be kept for up to five years after the end of the project and will then be destroyed.

You can withdraw from the study before December 15, 2011 by emailing or phoning me. You will find all contact details at the end of the questionnaire.

It will take 15 to 20 minutes to complete this survey. Please click on "I agree to participate" to allow me to use your data.

Thank you in advance!

Friederike Tegge

- **O** <u>I agree to participate</u>
- **O** I do **not** agree to participate

If you would like a summary of this study, please tick the box below and provide an e-mail address.

□ I would like to receive a summary and here is my e-mail address: _____

A Quick Overview of the Questionnaire:

The following questionnaire will ask you about your views on the use of songs in the language classroom, your actual teaching practice involving songs and some background information. I am keenly interested in your opinion and your experiences, regardless of whether you use songs or not.

One important note: If you consider yourself to be a language teacher but you are not currently teaching, please think of the last time you were actively involved in language instruction and answer the questions accordingly.

Questions 1 to 23 will ask you about your views on the use of songs in the classroom and your actual teaching practice involving songs.

Questions 24 to 31 will ask you to provide some information about yourself, such as your age, gender and teaching experience.

Questions 32 to 37 will ask you to provide some information about your musical interests and training. Depending on your answer choices, you might skip a few questions. Some Technical Information: If a word is highlighted in blue, you can hover over it with the cursor and it will give you some additional information. Try it. Different questions provide varying answer options (e.g. a drop-down menu, a text box, or multiple-choice answers). Once you have given your answer, click on the button with the arrows aiming forward (>>) . If you would like to change your answer to a previous question, go back to it by using the back button with the arrows pointing backwards (<<). If you accidentally failed to choose any answer, the programme will remind you by giving you instructions to answer in a big orange box. In that case, there is no need to click on the back button. Simply choose an answer and continue. Let's get cracking!

I. Your opinion and teaching practice: Songs in the Classroom

1. What language do you teach?

If you teach more than one language, please choose one and answer all questions thinking of that particular language.

- **O** Arabic
- O Chinese (Mandarin)
- O ...
- O Vietnamese
- O Other, please specify: _____

220

2. What is your opinion on the following sentence?

"Songs are a useful tool in the language classroom to foster language acquisition."

Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
0	0	0	0	0

3. What is your opinion on the following sentence when considering the different proficiency levels?

"Songs are a useful tool in the language classroom to foster language acquisition."

	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
Absolute beginners	•	О	0	0	0
Continuing beginners/ elementary beginners	0	0	0	0	0
Low- intermediate learners	0	О	o	0	О
High- intermediate learners	0	0	0	0	0
Low-advanced learners	O	О	o	O	О
High-advanced learners (near- native)	0	0	0	0	0

4. What is your opinion on the following sentence when considering the different class types?

	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
Second language courses for general, everyday purposes (e.g. ESL, DaZ)	0	O	0	O	0
Foreign language courses for general, everyday purposes (e.g. EFL, DaF)	O	O	O	O	0
Language courses for academic purposes, i.e. for study purposes in higher education (e.g. EAP)	0	O	O	O	0
Language courses for specific purposes, i.e. for specific professional or academic fields such as business, tourism or law (e.g. ESP)	O	O	O	O	O
Conversation classes	Ο	О	Ο	О	0

"Songs are a useful tool in the language classroom to foster language acquisition."

5. Do you use songs in the language classroom?

- O Yes
- O No

5/2. Why don't you use songs in the language classroom?

- □ I can't find any suitable songs.
- □ I don't have the necessary technical equipment.
- □ The technical equipment is unreliable.
- □ I don't think songs are useful for language learning in the classroom.
- □ I can't fit songs in with the official curriculum.
- □ I tried using songs but my students did not respond well.
- □ Using songs means spending too much time on too little learning outcome.
- Other, please explain: _____

6. Do you use songs in your classes when teaching the following proficiency levels?

	Yes	No	Not Applicable (I don't teach this level)
Absolute beginners	0	O	O
Continuing beginners/ elementary beginners	0	•	•
Low-intermediate learners	0	0	0
High-intermediate learners	0	0	0
Low-advanced learners	0	0	0
High-advanced learners (near native)	0	0	0

223

7. Do you use songs in your classes when teaching the following class types?

	Yes	No	Not Applicable (I don't teach this class type)
Second language courses for general, everyday purposes (e.g. ESL, DaZ)	0	•	0
Foreign language courses for general, everyday purposes (e.g. EFL, DaF)	•	•	О
Language courses for academic purposes, i.e. for study purposes in higher education (e.g. EAP)	0	0	0
Language courses for specific purposes, i.e. for specific professional or academic fields such as business, tourism or law (e.g. ESP)	0	0	0
Conversation classes	0	Ο	0

8. For what purposes do you use songs in the language classroom? (Multiple answers possible)

- **D** To create a relaxing atmosphere
- □ To motivate students with an enjoyable activity
- □ To give students the opportunity to produce language without feeling observed
- □ To teach authentic language and culture
- □ To practise listening comprehension
- □ To practise fluency in speaking i.e. producing spoken language
- **D** To teach pronunciation and prosody
- □ To introduce new vocabulary
- □ To practise familiar vocabulary
- **D** To teach multi-word units, i.e. idioms and phrasal language
- □ To introduce new or practise familiar grammatical items
- □ As a prompt for spoken interaction, e.g. a class or pair discussion
- □ As a prompt for a writing assignment, e.g. an essay, poem or letter
- □ To accommodate different learning styles, e.g. auditive, kin-aesthetic and musical learning styles

Other, please explain: _____

224

9. A colleague asks you whether you can recommend any songs she could use in her language class.

What 3 songs do you recommend (artist & title)?

For what proficiency level do you recommend them?

For what purpose do you recommend them?

And why do you recommend these particular songs?

Please fill in what you can, even if you do not have all the information. For example, if you don't know the title you can provide the first line or first few words.

Example:

	Artist	Title	Proficiency Level	Purpose	Why this particular song?
Song 1	Supertramp	Logical song	Upper intermediate learners	To introduce a number of adjectives/adverbs and to teach about typical suffixes of adjectives (-ful, - able etc.)	The tune is catchy and many students can relate to the theme of having to conform to society and "be like everybody else"

If you think of more songs later, feel free to send me an e-mail even after you've completed this survey. My contact details can also be found at the end of the questionnaire.

	Artist	Title	Proficiency Level	Purpose	Why this particular song?
Song 1					
Song 2					
Song 3					



10. What activities - if any - do you use in the classroom that are directly related to the song you have used? (Multiple answers possible)

- □ True/false statements
- □ Focus or comprehension questions
- □ Ordering activity, i.e. putting verses, lines, parts of lines or words in order
- □ Cloze/gap-fill activity
- Dictation (full or partial text)
- □ Sing-along
- Writing activity, e.g. composing an additional verse, writing a letter or writing an essay about the song's topic
- Discussion
- Other activities, please explain: ______

225

11. On average how many times do you have students sing along with the songs during one lesson?

- O Never
- O Once
- O Twice
- **O** Three times
- **O** More than three times

12. Do you yourself sing along with your students?

- **O** Always
- O Often
- **O** Sometimes
- O Never
- **O** My students don't sing

13. How comfortable do you feel singing along with your students?

- **O** Very comfortable
- $\mathbf{O} \ \ \text{Comfortable}$
- **O** Neither comfortable nor uncomfortable
- **O** Uncomfortable
- **O** Very uncomfortable
- **O** I don't sing

14. How comfortable do you feel singing by yourself in front of your students?

- **O** Very comfortable
- **O** Comfortable
- **O** Neither comfortable nor uncomfortable
- **O** Uncomfortable
- **O** Very uncomfortable
- **O** I don't sing

15. On average how many times do you play a song during one lesson (including the sing-alongs)?

- O Once
- O Twice
- **O** Three times
- Four times
- **O** More than four times

16. Do you replay a song you have used in class in a later lesson?

- **O** Always
- O Often
- **O** Sometimes
- O Never

17. How much class-time on average do you spend on a song and its directly related activities?

- **O** 5 minutes or less
- **O** Between 5 and 15 minutes
- **O** Between 15 and 30 minutes
- **O** Between 30 and 50 minutes
- **O** More than 50 minutes

18. How would you describe the learners' general response to the use of songs in the classroom?

- **O** Very positive
- **O** Positive
- O Neutral
- **O** Negative
- **O** Very negative

227

19. What problems - if any - do you have when using songs in the classroom? (Multiple answers possible)

- □ I cannot find any suitable songs
- □ We do not have the appropriate equipment
- □ The equipment is unreliable
- □ The noise bothers other classes/teachers
- □ The use of songs conflicts with the curriculum
- □ Too time-consuming
- □ Problems with learner-discipline
- **D** The learners do not like to sing
- □ The learners do not consider songs to be adequate for effective learning
- Other problems, please explain: _____

Ib. Your opinion and teaching practice: Songs and Vocabulary

20. What is your opinion on the following sentence?

"In general, songs are useful for introducing new vocabulary in the classroom."

Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
0	0	0	0	0

20/2. Do you use songs for introducing new vocabulary to your students in the classroom?

O Yes

O No

21. What is your opinion on the following sentence?

"In general, songs are useful for <u>practising</u> familiar vocabulary in the classroom."

Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
0	0	0	0	0

21/2. Do you use songs for practising familiar vocabulary with your students in the classroom?

- O Yes
- O No

22. Which of the following aspects of songs - if any - make songs suitable for vocabulary learning? *(Multiple answers possible)*

- **D** The rhythm
- The melody
- □ The possibility to sing along
- **D** The rhyme
- □ The fact that one can listen to a song several times without getting bored
- □ The fact that one can sing a song several times without getting bored
- □ The fact that learners are likely to encounter a song we used in class outside of class
- □ The fact that songs usually contain authentic language
- Other aspects, please explain: _____

II. A bit about you ...

- 23. What is your gender?
- O Male
- ${\mathbf O}$ Female

24. How old are you? (Please type your age in the box below, and remember: This survey is absolutely anonymous!)

25. What is your nationality? (Please choose from the drop-down menu below)

O Afghan

O ...

26. What country are you currently teaching in? (Please choose from the drop-down menu below)

O AfghanistanO ...

26/2. In what country have you taught the most during your teaching career? (Please choose from the drop-down menu below)

(same list of countries as in 26)

27. At what type(s) of institution are you currently teaching? (Multiple answers possible)

- Primary school
- □ Secondary school
- Dest-secondary vocational training school (e.g. trade school)
- □ University, tertiary college, polytechnic
- □ Adult education centre, continuing education institution
- Public or private language school
- □ Self-employed private teaching
- Other, please specify: _____

27/2. At what type(s) of institution did you teach before? (Multiple answers possible)

- Primary school (1)
- □ Secondary school (2)
- Dest-secondary vocational training school (e.g. trade school) (3)
- □ University, tertiary college, polytechnic (4)
- □ Adult education centre, continuing education institution (5)
- Public or private language school (6)
- □ Self-employed private teaching (7)
- Other, please specify: _____

28. How many years of experience as a language teacher do you have?

- **O** Less than 1 year
- **O** 1 to 2 years
- **O** 3 to 5 years
- O 6 to 10 years
- **O** More than 11 years

29. What proficiency levels are you currently teaching or have you previously taught? (Multiple answers possible)

- □ Absolute beginners
- □ Continuing beginners/elementary beginners
- □ Lower intermediate learners
- **Upper intermediate learners**
- Lower advanced learners
- □ Higher advanced learners (near-native)

30. What course types are you currently teaching or have you previously taught?

- □ Second language courses for general, everyday purposes (e.g. ESL, DaZ)
- □ Foreign language courses for general, everyday purposes (e.g. EFL, DaF)
- □ Language course for academic purposes, i.e. for study purposes in higher education (e.g. EAP)

- □ Language course for specific purposes, i.e. for specific professional or academic fields such as business, tourism or law (e.g. ESP)
- Conversation class
- Other, please specify: _____

231

31. What is the highest degree of training in language teaching you currently hold?

- **O** No degree involving training in language teaching
- **O** A language teaching certificate from a public or private institution (e.g. CELTA)
- **O** An undergraduate degree (e.g. a Bachelor's degree)
- **O** A postgraduate degree (e.g. a Master's degree or a PhD)
- O Other, please specify: _____

III. Your musical interests and training

32. How important is music in your personal life?

- **O** Very Important
- **O** Important
- **O** Neither Important nor Unimportant
- $\mathbf{O} \quad Unimportant$

33. How often do you intentionally listen to the following kinds of music? (This does not include involuntary listening to background music in supermarkets, the gym etc.)

	Very Often	Often	Sometimes	Never
Instrumental, non- vocal music	0	О	О	О
vocal music	O	Ο	O	Ο

34. Have you ever done the following activities and if yes for how long?

	Never	Less than 1 year	1 to 2 years	3 to 5 years	6 to 10 years	More than 11 years
Sing in a music ensemble (e.g. a choir, band etc.)	0	0	0	0	0	0
Play an instrument	O	o	O	o	O	O

35. Can you read music?

O Yes

O No

36. When you yourself were a language learner, did your language teacher use songs in the classroom?

- O Yes
- O No

36/2. Overall, how did you as a language learner experience the use of songs in the language classroom?

Please finish the following sentence with the appropriate choice:

"The experience was overall ...

- **O** Very Positive
- **O** Positive
- **O** Neither Positive Nor Negative
- **O** Negative
- **O** Very Negative

37. If there is anything I left out regarding the use of songs in the language classroom and you would like to comment on, this is the place: _____

Thank you very much!



If you have further questions regarding the project, feel free to contact me or my supervisors Assoc Prof Frank Boers and Dr Averil Coxhead. Our contact details are: Friederike.Tegge@vuw.ac.nz +64 4 463 5233 extn 8029, Frank.Boers@vuw.ac.nz +64 4 463 6014, Averil.Coxhead@vuw.ac.nz +64 4 463 5604, School of Linguistics and Applied Language Studies, PO Box 600, Victoria University of Wellington, Wellington 6140 New Zealand.

This PhD project is being conducted in the School of Linguistics and Applied Language Studies at Victoria University of Wellington, New Zealand, and has been approved by the university's Human Ethics Committee.
Appendix A.2: Ethics approval for the teacher survey



MEMORANDUM

Phone 0-4-463 5676

Fax 0-4-463 5209

то	Friederike Tegge
COPY TO	Frank Boers & Averil Coxhead
FROM	Dr Allison Kirkman, Convener, Human Ethics Committee
DATE	21 January 2011
PAGES	1
SUBJECT	Ethics Approval: No 18192 Vocabulary learning through songs? – The effects of rhythm, rhyme, melody, and oral reproduction on the recollection of foreign / second language vocabulary

Thank you for your application for ethical approval, which has now been considered by the Standing Committee of the Human Ethics Committee.

Your application has been approved from the above date and this approval continues until 31 July 2013. If your data collection is not completed by this date you should apply to the Human Ethics Committee for an extension to this approval.

Best wishes with the research.

Allison Kirkman

Human Ethics Committee

Appendix B: Materials for Study 2 (the pedagogical song corpus)

- Appendix B.1: A list of ESL textbooks used as sources for the compilation of the song corpus
- Appendix B.2: A list of ESOL-teacher websites used as sources for the compilation of the song corpus
- Appendix B.3: A list of lexical vocables found in the song corpus
- Appendix B.4: A list of songs included in the pedagogical song corpus

Appendix B.1: A complete list of ESL textbooks used as sources for the compilation of the song corpus

- 1. Cutting Edge Starter (2002)
- 2. Cutting Edge Advanced (2003)
- 3. English File 1 (2001)
- 4. English File Intermediate (2001)
- 5. English Matters Advanced (2000)
- 6. face2face elementary (2005)
- 7. face2face Pre-Intermediate (2005)
- 8. face2face Intermediate (2006)
- 9. Inside Out Elementary (2003)
- 10. Inside Out Pre-Intermediate (2002)
- 11. Inside Out Intermediate (2000)
- 12. Inside Out Upper Intermediate (2001)
- 13. Inside Out Upper Advanced, Teacher's Book (2001)
- 14. New Cutting Edge Elementary (2005)
- 15. New Cutting Edge Pre-Intermediate (2005)
- 16. New Cutting Edge Intermediate (2005)
- 17. New Frame Elementary (2008)
- 18. New Headway Elementary (2006)
- 19. New Headway Pre-Intermediate (2007)
- 20. New Headway Intermediate (2009)
- 21. New Headway Upper Intermediate (2005)
- 22. New Matrix Intermediate (2006)
- 23. Snapshot Intermediate (2000)
- 24. World View 4 (2002)

Appendix B.2: A complete list of ESOL-teacher websites used as sources for the compilation of the song corpus

- 1. www.busyteacher.org
- 2. www.eslvideo.com
- 3. www.eslcafe.com
- 4. www.esl-galaxy.com/music.htm
- http://englishwithjeniffer.wordpress.com/2010/07/21/tell-us-all-about-it-more-useof-good-lyrics/
- http://www.slideshare.net/grace6671/pop-songs-in-the-englishclass?src=related_normal&rel=70185
- 7. www.teachingenglishgames.com/Articles/Using_ESL_Songs.htm

Appendix B.3: A complete list of lexical vocables found in the song corpus

АААН	-DYF	ОН
AAH	FF	ОНН
AAOW	FFH	ОННН
AH	FH	оно
АНН	FY	OOH
AHA	FAH	0000
AI	НА	000H
Αυμμ	НАА	0000H
AW	НААА	00
AY	HADY	OY
BA	НАН	RONN
BADABUP	НАҮР	ROOTY
BADADUP	HDAH	RUMPA
BISMILLAH	HEE	SHA
BOOHM	HEEY	SHINGA
BOM	HEH	SHOOBEE
BOMPA	HEI	SHOOBEDOO
BOP	HFLA	TOOT
BRAH	HELOA	TOOTS
BUBBA	HIPP	TRA
BUMM	НМ	TUMM
СНА	НММ	TUMMS
CHEE	НМММ	UAH
СНІСКА	НО	UH
CHING	НОО	UUH
DA		
DADN	нимм	UHUH
DAH	IDDI F	UM
DAI	ин ИН	WA
DAMM	IA	WAH
DF	JIGGLE-A-MESA-CARA	WANG
DEE	JIMINY	WFF
DFFDLF	JUDA	WFFH
DIB		WEEMA
DIBBA	LALA	WFH
DICKY-	LAMM	WFH
DIDDI	LE	WHOA
DIDDL	LING	WHOAH
DIDLE	LINGA	WO
DILLER	МАМВАН	WOAH
DILLIE	MEE	WOH
DJANGO	MM	WOO
DOO	MMM	WOOH
DOOBEE	MMMM	WOOHOO
DOODLE	NA	WOOHOO
DOODN	OBLADA	YAH
DOOTN	OBLADI	YAI
DUM	00	YEAH
DUMM	000	YI

Performer	Title
10 cc	I'm not in love
2 Unlimited	Info Superhighway
3oh!3, Katy Perry	Starstrukk
АВВА	Dancing Queen
АВВА	Fernando
ABBA	Happy New Year
ABBA	I have a dream
ABBA	Money Money
ABBA	The day before you came
ABBA	The winner takes it all
Ace of Base	All that she wants
Adele	Rolling in the deep
Adele	Set fire to the rain
Adele	Hometown glory - London riots version
Aerosmith	I don't want to miss a thing
Akon	Right now (Na, na, na)
Alabama	Angels among us
Alanis Morissette	Hand in my pocket
Alanis Morissette	Ironic
Alanis Morissette	That I would be good
Alanis Morissette	You owe me nothing in return
Alanis Morissette	Perfect
Alesha Dixon	The boy does nothing
Alexander Rybak	Foolin'
Alicia Keys	Empire state of mind (Part II)
Alicia Keys	Streets of New York
All-American Rejects, the	Gives you hell
All Saints	Never ever
Alphanaut	Never been to Athens
Amy Winehouse	Rehab
Amy Winehouse	You know I'm no good
Amy Winehouse	Back to black
Anastacia	Left outside alone
Anita Renfroe	Total Momsense
Aqua	Barbie girl
Avett Brothers, the	I and Love and You
Avril Lavigne	Complicated
Avril Lavigne	I'm with you
Avril Lavigne	Innocence
Avril Lavigne	sk8ter boi
Avril Lavigne	Things I'll never say
Avril Lavigne	When you're gone

Appendix B.4: List of songs included in the pedagogical song corpus

Backstreet Boys, the	As long as you love me
Backstreet Boys, the	I want it that way
Backstreet Boys, the	Shape of my heart
Bananarama	Love in the first degree
Bangles, the	Manic Monday
Barenaked Ladies	If I had a million dollars
Barry Manilow	Can't smile without you
Barbra Streisand	Who's afraid of the big bad wolf?
Baz Luhrmann	Everybody's Free to Wear Sunscreen
Beach Boys, the	California Dreamin'
Beach Boys, the	Fun Fun Fun
Beatles, the	A day in the life
Beatles, the	A hard day's night
Beatles, the	All together now
Beatles, the	All You Need Is Love
Beatles, the	And I love her
Beatles, the	Because
Beatles, the	Eight days a week
Beatles, the	Eleanor Rigby
Beatles, the	Hello, Goodbye
Beatles, the	Help
Beatles, the	Here comes the sun
Beatles, the	Hey Jude
Beatles, the	I want to hold your hand
Beatles, the	If I fell
Beatles, the	Let it be
Beatles, the	Money (that's what I want)
Beatles, the	Obladi Oblada
Beatles, the	She loves you
Beatles, the	She's leaving home
Beatles, the	Till there was you
Beatles, the	When I'm 64
Beatles, the	Yesterday
Beatles, the	Yellow submarine
Beatles, the	You're going to lose that girl
Bee Gees	Night Fever
Ben Folds	Gracie
Ben E. King	Stand by me
Bette Midler	From a distance
Bette Midler	I'm a woman
Beyonce	Broken-hearted girl
Beyonce	If I were a boy
Beyonce	Single ladies
Bic Runga	Drive

Bic Runga	Get Some Sleep
Bic Runga	Listening for the weather
Bic Runga	Sway
Billy Joel	Just the way you are
Billy Joel	Piano man
Bill Withers	Lean on me
BJ Thomas	Raindrops keep falling on my head
Black Eyed Peas, the	I gotta feeling
Black Eyed Peas, the	Where is the love
Blink 182	First date
Blondie	The tide is high
Blue	l can
Blue	If you come back
Blur	Country house
Blue Rodeo, Sarah McLachlan	Dark angel
Bob Dylan	Blowin' in the Wind
Bob Dylan	Ballad of Hollis Brown
Bob Dylan	Dignity
Bob Dylan	Standing in the doorway
Bob Marley	One Love
Bob Marley	Three little birds
Bobby McFerrin	Don't worry, be happy
Boney M.	Calendar song
Boney M.	Let it all be music
Bon Jovi	Bad medicine
Bon Jovi	Wild in the streets
Bon Jovi	Ugly
Boyz 2 Men	A song for mama
Boyzone	No matter what
Brand New Heavies	You've got a friend
Brandy	Have you ever
Brad Paisley, Dolly Parton	When I get where I am going
Brian Hyland	Itsy, bitsy, teenie, weenie, yellow polka dot bikini
Brian McFadden	Real to me
Britney Spears	Hold it against me
Britney Spears	Lucky
Britney Spears	My only wish this year
Bruno Mars	Just the way you are
Bruno Mars	Marry you
Bruno Mars	Talking to the moon
Bruno Mars	The lazy song
Bruno Mars	Grenade
Bruce Springsteen	Into the fire
Bruce Springsteen	My hometown

Bruce Springsteen	Talk to me
Bryan Adams	Everything I do
Bryan Adams	Have you ever really loved a woman
Bryan Adams	On a day like today
Burt Bacharach	Do you know the way to San Jose?
Dionne Warwick	Trains and boats and planes
Burt Bacharach, Hal David	24 hours from Tulsa
Burt Bacharach, Hal David, Jackie DeShannon	What the world needs now is love
Caro Emerald	A night like this
Carpenters, the	Please Mr. Postman
Carpenters, the	We've only just begun
Carpenters, the	Yesterday once more
Carpenters, the	Jambalaya
Carly Simon	Nobody does it better - James Bond Theme
Carly Simon	You're so vain
Cascada	Every time we touch
Cascada	Last Christmas
Cat Stevens	Father and son
Cat Stevens	Old schoolyard
Cat Stevens	Where do the children play?
Celine Dion	Because you loved me
Celine Dion, Brahms	Brahms' Lullaby
Celine Dion	My heart will go on
Celine Dion	The power of love
Celine Dion	A new day has come
Celine Dion	At seventeen
Cher	Believe
Cher	If I could turn back time
Chordettes, the	Mr. Sandman
Christina Aguilera	Beautiful
Christina Aguilera	The voice within
Chris Brown	Crawl
Chris Brown, Justin Bieber	Next 2 you
Tony Christie, Peter Kay	Is this the way to Amarillo
Chris Madin	Dressed in white
Christy Moore	City of Chicago
Chris Rea	Driving home for Christmas
Clash, the	I'm not down
Cliff Richard	Summer holiday
Colbie Caillat	Midnight bottle
Coldplay	Christmas lights
Coldplay	Fix You
Coldplay	Viva la vida
Coldplay	What if

Cole Porter	Miss Otis Regrets
Corrs, the	Don't say you love me
Corrs, the	What can I do
Counting crows	Accidentally in love
Cowboy Junkies	If you were the woman and I was the man
Craig David	Walking away
Crosby, Stills, Nash and Young	Our house
Crystals, the	And then he kissed me
Crystals, the	Da Do Ron Ron (When he walked me home)
Cure, the	Boys don't cry
Cure, the	Friday I'm in Love
Cure, the	Killing an Arab
Cyndi Lauper	True colors
Damien Rice	Cold water
Damien Rice	The blower's daughter
Daniel Powter	Bad Day
David Bowie	Space Oddity
David Bowie	Fashion
Dave Dobbyn	Welcome home
David Guetta	When love takes over
Des'Ree	You gotta be
Diana Krall	The boy from Ipanema
Diana Krall	Pick yourself up
Diana Krall	It's wonderful
Diana Krall	So nice
Dido	Here with me
Dido	Mary's in India
Dido	Thank you
Dido	White flag
Dion and the Belmonts	Runaround Sue
Dionne Warwick	Always something there to remind me
Dionne Warwick	Walk on by
Disney	A whole new world
Disney	Colors of the wind
Disney	The lion sleeps tonight
Dixie Cups	Chapel of love
DJ Ötzi	hey baby
Dolly Parton	9 to 5
Dolly Parton	Coat of many colors
Don McLean	Vincent
Donovan	Universal soldier
Don Williams	Long black veil
Doors, the	People are strange
Duffy	Warwick Avenue

Duran Duran	Is there something I should know
Eagles, the	James Dean
Earth Mama	This River
East 17	Stay another day
Educational - Schoolhouse Rock	Conjunction Junction
Educational - Various	ABC/The Alphabet
Educational - Animaniacs	Nations of the World
Educational - Saxman 45/the Teacher and the	
Rockbots	Where will I go?
Educational - Sesame street	One fine face
Edwin Starr	War
Eiffel 65	Blue
Eleanor McEvoy	Sophie
Ella Fitzgerald	Santa Clause got stuck in my chimney
Ell and Nikki	Running scared
Elton John	Believe
Elton John	Candle in the wind
Elton John	Daniel
Elton John	Sad songs (say so much)
Elton John	Your song
Elvis Presley	Hound dog
Elvis Presley	Love Me Tender
Elvis Presley	Return to sender
Elvis Presley	His latest flame
Elvis Presley	Suspicious minds
Emilia	Big big world
Eminem, Rihanna	Love the way you lie
Erasure	Oh L'amour
Eric Bogle	I hate wogs
Eric Clapton	Tears in heaven
Eric Clapton	Wonderful tonight
Etta James	At last
Eva Cassidy	Somewhere over the rainbow
Eva Cassidy	Tennessee waltz
Everly Brothers	All I have to do is dream
Everything but the girl	Missing
Fiddler on the Roof	If I were a rich man
Flaming Lips, the	Fight Test
Fools Garden	Lemon Tree
Frank Sinatra	It happened in Monterey
Frank Sinatra	My way
Frank Sinatra	New York, New York
Gabrielle	Out of reach
Garbage	# 1 Crush
Gareth Gates	Too serious too soon

Gene Autry	Rudolph
Genesis	No son of mine
Gene Kelly	Singing in the rain
Gerry and the Pacemakers	You'll never walk alone
Gerry Rafferty	Right Down the Line
Girls Aloud	I'll stand by you
Glee	Billionaire
Glee, Lily Allen	Smile
Global Deejays	Sound of San Francisco
Gloria Estefan	Words get in the way
Gloria Gaynor	I will survive
Gordon Lightfoot	If you could read my mind
Grease	Summer Nights
Green Day	Boulevard of broken dreams
Green Day	I fought the law
Guns n Roses	Don't Cry
Guns n Roses	Patience
Guns n Roses	Sweet child o' mine
Harry Chapin	Cats in the Cradle
High School Musical	What time is it?
Hilary Duff	Someone's watching over me
Hollies, the	Bus Stop
Hoobastank	The reason
Hugh Grant, Haley Bennett	Way Back Into Love
Human League, the	Don't you want me
Ian Dury and the Blockheads	I believe
Ingrid Michaelson	Keep Breathing
lyaz	Replay
Jackson Browne	Lives in the Balance
Jackson Five, the	Santa Clause is coming tonight
Jackson Five, the	I saw mommy kissing Santa Claus
Jack Johnson	Better together
Jack Johnson	The 3 R's
Jack Johnson	We're going to be friends
Jack Johnson	With my own two hands
James Blunt	High
James Blunt	Goodbye my lover
James Blunt	No Bravery
James Blunt	Same mistake
James Blunt	You're Beautiful
James Taylor	Fire and rain
Janis lan	Matthew
Jason Mraz	I'm Yours
Jason Mraz	Life is wonderful

Jason Mraz	Lucky
Jason Mraz	The sunshine song
Jeff Buckley	Hallelujah
Jeffery Lewis	Life
Jenny Lewis	Barking at the moon
Jennifer Lopez	Do you know where you're going to?
Jennifer Lopez	If you had my love
Jessie J	Price Tag
Jet	Are you gonna be my girl?
Jevetta Steele	Calling You
Jimmy Cliff	I can see clearly now
Jimmy Cliff	You can get it if you really want
Jim Cuddy	Everyone watched the wedding
Jim Rees	Amazing Grace
Jim Reeves	Jingle Bells
Joan Armatrading	Me Myself I
Joan Osborne	One of us
Joey Scarbury	Believe it or not
Johnny Cash	A boy named Sue
Johnny Cash	Folsom prison blues
John Denver	Country roads
John Denver	Leaving on a jet plane
John Denver	Rocky Mountain High
John Denver	Sunshine on my shoulders
John Lennon	Imagine
John Martyn	Couldn't love you more
John Mayer	Waiting on the world to change
John Prine	Please don't bury me
John Paul Young	Love is in the air
Joseph M. Scriven	What a friend we have in Jesus
J.P. Taylor	Rainforest Song
Judy Garland	How about you (Babes on Broadway)
Julie Andrews	Getting to know you
Justin Bieber	Baby
Justin Bieber	One time
Justin Bieber	Somebody to love
Just Jack	The day I died
Justin Timberlake	What goes around
Katie Melua	If you were a sailboat
Katy Perry	E.T. (Futuristic Lover)
Katy Perry	Firework
Katy Perry	Last Friday night (T.G.I.F.)
Katy Perry	Hot N cold
Keane	Somewhere only we know

Keb Mo	I'm Amazing
Kelly Clarkson	Because of you
Kelly Clarkson	Breakaway
Kenny Rogers	You Decorated My Life
Killers, the	Mr. Brightside
Kings of Convenience	Misread
Kinks, the	Dedicated follower of fashion
K'naan	Wavin' flag
Kraftwerk	The telephone call
Lady Gaga	Born this way
Lady Gaga	Just dance
Lara Fabian	Adagio
Led Zeppelin	Stairway to heaven
Lemonheads, the	Being around
Lenka	Everything at once
Lenny Kravitz	I'll be waiting
Leonard Cohen	Hey, that's no way to say goodbye
Leonard Cohen	I'm your man
Lesley Gore	It's my party
Linkin Park	Leave out all the rest
Linkin Park	Numb
Linkin Park	What I've done
Linda Rondstadt	What'll I do
Lisa Stansfield	All around the world
Loretta Lynn	Success
Louis Armstrong	What a wonderful world
Louis Armstrong	Summertime
Lou Bega	Mambo No. 5
Lou Reed	Dirty Boulevard
Lou Reed	Perfect Day
Luther Vandross	Dance with my father
Lyle Lovett	Give back my heart
m2m	The day you went away
Madness	My girl
Madness	Our house
Madonna	American pie
Madonna	Frozen
Madonna	Hung up
Madonna	I deserve it
Madonna	La isla bonita
Madonna	Material girl
Madonna	Nothing really matters
Madonna	This used to be my playground
Madonna	Lucky star

Madonna	Holiday
Maher Zain	Open your eyes
Malaysian Artistes For Unity	Here In my Home
Mamas and the Papas, the	Dancing in the street
Mandy Moore	Cry
Mariah Carey	All I want for Christmas is you
Mariah Carey	Hero
Mariah Carey	I want to know what love is
Mariah Carey	Shake it off
Marianne Faithful	This little bird
Marvin Gaye	What's going on?
Marcia Griffiths	The Electric Boogie
Martina McBride	Anyway
Martina McBride	Concrete Angel
Martina McBride	Do you hear what I hear
Marilyn Monroe	Happy birthday, Mr. President
Maroon 5	This love
Maroon 5	Won't go home without you
Mia	Paper planes
Michael Bolton	I said I loved you but I lied
Michael Buble	Everything
Michael Buble	Haven't met you yet
Michael Buble	Home
Michael Franti, Spearhead	Bomb the World
Michael Jackson	Beat it
Michael Jackson	Billie Jean
Michael Jackson	Earth song
Michael Jackson	Heal the world
Michael Jackson	Hold my hand
Michael Jackson	Man in the Mirror
Michael Jackson	Thriller
Michael Jackson	We are the World
Michael Learns To Rock	Take me to your heart
Michael W. Smith	Great is the Lord
Mika	Rain
Miley Cyrus	Seven things
Miley Cyrus	Life's a climb
Monty Python	Always look on the bright side of life
Muse	Can't take my eyes off you
Muse	Uprising
Mutton Birds, the	April
Mutton Birds, the	Come around
Mutton Birds, the	Wellington
ΝΑΗΡΙ	Do They Know it's Halloween? (A Benefit for UNICEF)

Nancy Sinatra	These boots are made for walking
Natasha Bedingfield	Unwritten
Neil Young	Heart of gold
Nelly Furtado	I'm like a bird
Ne-Yo	So sick
Nickelback	If today was your last day
Nickelback	Photograph
Nickelback	Rock Star
Nickelback	If everyone cared
Nina Simone	Beautiful Land
Nirvana	Smells Like Teen Spirit
Noa	Blue touches blue
Noa	Lady Night
Noa	Mark of Cain
Nora Jones	Come away with me
Oasis	Live forever
Oasis, Coca Cola	Whatever
Oi Va Voi	Foggy Day
OneRepublic	All the right moves
OneRepublic	Apologize
OneRepublic	Marchin' on
OneRepublic	Secrets
Orianthi	According to You
Owl City	Fireflies
Shihad, Pacifier	Only time
Paper Lace	Billy don't be a hero
Paramore	The only exception
Patsy Cline	She's got you
Paula Abdul	Opposites Attract
Paul Kelly	Leaps and Bounds
Paul Williams; the Carpenters	Rainy days and Mondays
Pearl Jam	Wishlist
Peggy Seeger, Ewan MacColl	The first time I ever saw your face
Pet Shop Boys	Always on my mind
Pete Seeger	If I had a hammer
Phil Collins	Another day in paradise
Pink Floyd	Mother
Pink Floyd	Another brick in the wall Part II
Pink	Dear Mr President
Pink	Don't let me get me
Pink	Family portrait
Pink	Just like a pill
Pink	Raise your glass
Pink	Stupid Girls

Pirates of the Caribbean (Disney)	Yo ho - a pirate's life for me
Platters, the	Only you
Pogues, the	Fairy-tale of New York
Police, the	Every breath you take (< Sting)
Queen	Bohemian rhapsody
Queen	Don't stop me now
Queen	I'm going slightly mad
Queen	Radio Ga Ga
Queen	The show must go on
Queen	We are the champions
Queen	You're my best friend
Radiohead	Fake plastic trees
Rambling Wheels, the	Wild
Rasmus, the	In the shadows
Ray Peterson	Tell Laura I love her
Reamonn	Tonight
Rebecca Black	Friday
Rembrandts, the	I'll be there for you
Righteous Brothers, the	You've lost that lovin' feelin'
Rihanna	Umbrella
R. Kelly	I believe I can fly
Robert Ramirez	Sick of love
Robbie Williams	Feel
Robbie Williams	Angels
Robbie Williams	That's life
Robbie Williams, Garry Barlow	Shame
Robbie Williams	Beyond the sea
Rod Stewart	Sailing
Rolling Stones, the	Beast of burden
Rolling Stones, the	Satisfaction
Roxette	Fireworks
Roy Orbison	You got it
Royal Teens, the	Short shorts
Ruru Karaitiana	Blue Smoke
Sade	Somebody already broke my heart
Sam Cooke	Wonderful World
Sandra	(I'll never be) Maria Madgalena
Santa Esmeralda	You are my everything
Santana, Rob Thomas	Smooth
Sarah McLachlan	I will remember you
Sarah McLachlan	Ordinary miracle
Savage Garden	Truly madly deeply
Scorpions	Wind of change
Script, the	The man who can't be moved

Shakira	She wolf
Shakira	Waka waka
Shakin' Stevens	Merry Christmas everyone
Shania Twain, Anne Murray	You needed me
Shania Twain, Bryan White	From this moment on
Shayne Ward	That's my goal
Sheryl Crow	My favorite mistake
Sick Puppies	All the same
Simon and Garfunkel	Bridge over troubled water
Simon and Garfunkel	El condor pasa
Simon and Garfunkel	Homeward bound
Simon and Garfunkel	Old friends
Simon and Garfunkel	Sound of silence
Sinead O'Connor	Nothing compares to you
Smash Mouth	I'm a believer
Smashproof	Brother
Smokey Robinson	Tears of a clown
Solid British Hat Band	Mister Monday
Sonny and Cher	l got you babe
Soulsister	The way to your heart
Sound of Music	Do-re-mi
Stacie Orrico	l promise
Stevie Ray Vaughan	Mary had a little lamb
Stevie Wonder	I just called to say I love you
Stevie Wonder	Living for the city
Stevie Wonder	You are the sunshine of my life
Sting	Englishman in New York
Sting	Fields of gold
Sting	Russians
Sting	Shape of my heart
Sting	Moon over Bourbon Street
Stylistics, the	You are everything
Sundays, the	Summertime
Supergrass	Alright
Supertramp	Logical song
Supremes, the	You Can't Hurry Love
Suzanne Vega	Luka
Suzanne Vega	Tom's Diner
Taio Cruz, Kylie Minogue	Taking us higher
Take That	Hello
Taylor Swift	Love story
Taylor Swift	Mine
Taylor Swift	You belong with me
Billie Holliday	Stormy Weather

Teddy Pendergrass	Wake up everybody
They Might Be Giants	Seven days of the week
Who, the	pinball wizard
Tim Finn, Herbs	Parihaka
Tito Nieves	I'll always love you
Tokio Hotel	Monsoon
Tom Tom Club	Wordy Rappinghood
Tony Bennett	I Left my heart in San Francisco
Toni Braxton	How could an angel break my heart
Tori Amos	Winter
Tracy Chapman	Behind the wall
Tracy Chapman	Fast car
Tracy Chapman	Baby can I hold you
Traditional	5 little monkeys
Traditional	Danny Boy
Traditional	Dark eyed sailor
Traditional	Down by the bay
Traditional	Eyes on the prize
Traditional	God defend New Zealand - Anthem NZ
Traditional	God Save the Queen - Anthem GB
Traditional	He's got the whole world in his hands
Traditional	Head, shoulders, knees and toes
Traditional	I like the flowers
Traditional	Old lady who swallowed a fly
Traditional	Old MacDonald
Traditional	Silent night
Traditional	Star Spangled Banner - Anthem USA
Traditional	The Eensy Weensy Spider
Traditional	The Wheels on the Bus
Traditional	Waltzing Matilda
Traditional	Mary had a little lamb
Train	Brick by brick
Train	Drops of Jupiter
Train	Hey soul sister
Travis	Turn
Travis	My eyes
Treat Her Right	I think she likes me
Turner, Ike & Tina	River deep, mountain high
U2	Sometimes you can't make it
U2	Stuck in a moment
U2	Trying to throw your arms around the world
U2	With or without you
U2	I still haven't found what I'm looking for
U2	One

Uncle Kracker	Smile
Uncle Monkey	Sunny days
Vamps	Love addict
Veronicas, the	Hook me up
Victor Mizzy and Orchestra	The Addams Family Theme Song
Wallflowers, the	One headlight
	Sailing to the other side (the Interislander
Warratahs, the	song)
Warratahs, the	Taranaki
Weezer	Island in the sun
Westlife	Flying without wings
Westlife	I lay my love on you
Westlife	My love
Westlife	Seasons in the sun
Westlife	The rose
Westlife	You raise me up
Wet Wet Wet	Love is all around
Wham!	Last Christmas
Woody Guthrie	So long, it's been good to know you
X Japan	Crucify my love
Zager/Evans	In the year 2525
Zain Bhikha	My mom is amazing

Appendix C: Materials for Study 3 (the intervention study)

- Appendix C.1: Ethics approval for the quasi-experimental intervention study
- Appendix C.2: Information sheet for participants in the intervention study
- Appendix C.3: Consent form for students participating in the intervention study
- Appendix C.3: Gap-fill activity (song version)
- Appendix C.4: Gap-fill activity (prose version)
- Appendix C.6: Free recall test
- Appendix C.7: Cued recall test
- Appendix C.8: Recognition test (song version)

Appendix C.1: Ethics approval for the quasi-experimental intervention study



MEMORANDUM

 Phone
 0-4-463 5676

 Fax
 0-4-463 5209

 Email
 Allison.kirkman@vuw.ac.nz

то	Friederike Tegge
COPY TO	Frank Boers Averil Coxhead
FROM	Dr Allison Kirkman, Convener, Human Ethics Committee
DATE	15 August 2012
PAGES	1
SUBJECT	Ethics Approval: 18192 Vocabulary Learning through songs? The effects of rhythm, rhyme, melody, and oral reproduction on the recollection of foreign/second language vocabulary

Thank you for your request to amend your ethics approval. This has now been considered and the request granted.

Your application has approval until 31 July 2013. If your data collection is not completed by this date you should apply to the Human Ethics Committee for an extension to this approval.

Best wishes with your research.

Allison Kirkman Human Ethics Committee

Appendix C.2: Information sheet for students participating in the intervention study

Information Sheet: "Activities in the English Language Classroom"

"Participate in linguistic research and be in to win an iPod shuffle!"

Kia ora!

My name is Friederike Tegge, and I am a PhD student at Victoria University of Wellington, New Zealand. I am carrying out research on learning English in the classroom. In particular, I am looking at classroom activities and how they affect language learning.

I would very much appreciate it if you could participate in my research project. Participation involves taking a short language proficiency test and then taking part in learning activities during one regular language lesson. These learning activities are similar to what you normally do in the language class, e.g. reading and listening to a short text, answering comprehension questions, discussing a topic with other students. Afterwards, I will evaluate how well these activities helped you learn. In addition, you will be asked to provide some information about yourself (e.g. your age, gender etc.). All of this will take about 90 minutes.

Your identity will not be disclosed. The data obtained will be stored in a locked or passwordprotected file and will only be accessed by me and my two supervisors, Dr. Frank Boers and Dr. Averil Coxhead. Neither your institution nor your teacher will have at any point access to your data and test results will not affect your class results in any way. All worksheets and questionnaires will be destroyed five years after the end of this research project. The data you provide may contribute to research publications, which will be publicly available.

I am happy to discuss further details and the exact purpose of this study with you once the research at this institution has finished.

If you agree to participate in this research and if you give permission to use your data in the way outlined above, please sign the attached consent form. If you would like to withdraw your data from the study at a later point, please notify me by email (address below) before January 31, 2013.

All students signing the consent form and participating in this research will automatically enter in a draw to win an iPod shuffle. The winner will be chosen randomly. If you would like to receive a summary report of this project, please tick the box on the attached consent form.

Thanks a lot!

Friederike Tegge

Friederike TeggeSchool of Linguistics and Applied Language StudiesVictoria University of WellingtonPO Box 600Wellington 6140New ZealandTel: + 64 4 463 5233 extn 8029Email: Friederike.Tegge@vuw.ac.nz AssociateProfessor Dr. Frank BoersTel: + 64 4 463 6014Email: Frank.Boers@vuw.ac.nz

Appendix C.3: Consent form for students participating in the intervention study

CONSENT TO PARTICIPATE IN RESEARCH

Friederike Tegge: "Activities in the English Language Classroom"

NAME (please print clearly):

I understand that this research involves taking a short language proficiency test and then taking part in several learning activities during my regular language lessons. The activities are similar to what I normally do in the language class. Afterwards Friederike will evaluate how well the activities helped me learn. I will also be asked to provide some information about myself in a questionnaire. All of this will take about 90 minutes.

I have had an opportunity to ask questions and have had them answered to my satisfaction.

I understand that the data I provide will be stored in a locked or password-protected file in the School of Linguistics and Applied Language Studies at Victoria University of Wellington in New Zealand. I also understand that my information will remain confidential. Information that may be used to identify me uniquely will not be stored together with my data.

I understand that I may withdraw any data or information I have provided before January 31, 2013.

I agree to take part in this research.

Signature		
0.0		

Date:

Please tick the box below if you would like a short summary of the study (and please provide an email address):

•	I would like to receive a summary of the study:		
---	---	--	--

My email address: ______

Appendix C.3: Gap-fill activity (song version)



SONG

My best friend gave me the best advice:

He said each day's a gift and not a _____ right.

Leave no stone _____, leave your fears behind,

and try to take the path less travelled by.

That first step you take is the longest stride.

If today was your last day

and tomorrow was too _____

could you ______ goodbye to yesterday?

Would you ______ each moment like your last?

Leave old pictures in the past?

_____ every dime you have?

If today was your last day.

Against the ______ should be a way of life.

What's worth the prize is always worth the fight.

Every second ______ 'cause there is no second try.

So live it like you'll never live it twice.

Don't take the free ride in your own life.

If today was your last day

and tomorrow was too _____,

could you ______ goodbye to yesterday?

Would you ______ each moment like your last?

Leave old pictures in the past?

_____ every dime you have?

Would you call old friends you never see?

_____ on memories?

Would you forgive your enemies?

And would you find that one you're dreaming of?

Swear up and down to God _____

that you'd finally fall _____ love?

If today was your last day.

If today was your last day,

would you ______ your mark by mending a broken heart?

You know, it is never too late to _____ for the stars

regardless _____ who you are.

So do whatever it takes,

'cause you can't _____ a moment in this life.

Let nothing stand in your way

'cause the ______ of time are never on your side.

If today was your last day

and tomorrow was too _____,

could you _____ goodbye to yesterday?

Appendix C.5: Gap-fill activity (prose version)



TEXT

My best friend gave me the best advice: He said each day's a gift and not a ______ right. Leave no stone ______, leave your fears behind. And try to take the path less travelled by. That first step you take is the longest stride.

Then he asked: If today was your last day and tomorrow was too ______, could you ______ goodbye to yesterday? Would you ______ each moment like your last, leave old pictures in the past, ______ every dime you have. If today was your last day?

Against the ______ should be a way of life. What's worth the prize is always worth the fight. Every second ______ 'cause there is no second try. So live it like you'll never live it twice. Don't take the free ride in your own life.

So: If today was your last day and tomorrow was too ______, could you ______ goodbye to yesterday? Would you ______ each moment like your last, leave old pictures in the past, ______ every dime you have? Would you call old friends you never see and ______ on memories? Would you forgive your enemies? Would you find that one you're dreaming of and swear up and down to God ______ that you'd finally fall ______ love. If today was your last day?

If today was your last day, would you ______ your mark by mending a broken heart? You know, regardless ______ who you are, it is never too late to ______ for the stars. So do whatever it takes, 'cause you can't ______ a moment in this life. Let nothing stand in your way 'cause the ______ of time are never on your side.

If today was your last day and tomorrow was too ______, could you ______ goodbye to yesterday?

Appendix C.6: Free recall test



Free Recall: Song

During the lesson you listened to a song. Do you remember the exact words of the song? Write down as much of the exact words as you remember. If you can, write them down in the order as they appeared in the song. If you are not sure about the order, just write them down in the order that you think looks most like the original lyrics. The first sentence has been done for you as an example.

My best friend gave me the best advice:				

Appendix C.7: Cued recall test



Cued Recall: Song

During the lesson with Friederike you listened to a song. Do you remember the exact words of the song? Below you can see the lyrics with some of the words missing and replaced by lines. Each line represents one missing word. So if you see three lines, then three words are missing there.

Please try to remember the exact lyrics and write them down, one word per one line. Shortened words like *don't* for *do not*, *'ll* for *will*, *'s* for *is* and *'d* for *would* are also counted as separate words. The first sentence has been done for you as an example.

1)	My best <u>f</u>	ríend	gave	me th	e <u>best</u>	advíce	_:	
2)	He said				a gift an	d not		
3)	Leave		the		, leave you	ur	behind, by.	
4)	That			you	is the	2	stride.	
5)		the				_ a way of	life.	
6)	What's wor	th the priz	e					
7)	 Every			'cause _			no second _	
8)	So		<u></u>		you'll never	live		·
9)	Don't	th	e		in			life.

263

10)	If today was			
	and tomorrow		/	
	you	goodbye	to yesterday?	
11)	Would you		like your	?
	Leave	in	past?	
	Donate	you h	ave?	
12)	Would you call		_ you	?
	on memories?			
13)	Would you	er	nemies?	
14)	And would you find		you	
	?			
	Swear up and down to _			
	that you'd		?	
15)	If today was		·	
16)	If today was		,	
	would you	your		a broken
	?			
17)	You, it is	too la	ate to	the
		who you	·	
18)	So do		,	
	'cause you can't			this life.
19)	Let		your	
	'cause the o	of time are never		·

264

Appendix C.8: Recognition test (song version)



<u>SONG</u>

During the lesson with Friederike you listened to a song. Do you remember the exact words of the song? Below you find sets of four sentences. Only one of them actually appeared word for word in the song. Try to find the sentence that matches exactly and word for word what you have heard in the song and tick that box (). The differences are <u>underlined</u>. But still read carefully: The differences between the four options might be quite small. If you are unsure, be spontaneous and go with your "gut-feeling".

Example:

A My **good** friend gave me **some good** advice:

B My **good** friend gave me **the best** advice:

- C My **best** friend gave me **some good** advice:
- My **<u>best</u>** friend gave me <u>the best</u> advice:
- A He said <u>every</u> day's a gift and not a given right.
- B He said <u>each</u> day's a gift and not a given right.
- C He said <u>every single</u> day's a gift and not a given right.
- D He said <u>each and every</u> day's a gift and not a given right.
- 2.

1.

- A ... leave your worries behind.
- B ... leave your <u>doubts</u> behind.
- C ... leave your <u>fears</u> behind.
- D ... leave your <u>troubles</u> behind.
- 3.
- A And try to <u>take</u> the <u>path</u> less travelled by.
- B And try to walk the path less travelled by.
- C And try to take the route less travelled by.
- D And try to walk the route less travelled by.

- 4.
- A That first step you take is the longest stride.

B That first step you take is <u>often</u> the longest stride.

C That first step you take is <u>always</u> the longest stride.

D That first step you take is <u>never</u> the longest stride.

- 5.
- A If today was the last day and tomorrow was too late, could you say goodbye to yesterday?

B If today was the last day and if tomorrow was too late, would you say goodbye to yesterday?

C If today was your last day and if tomorrow was too late, <u>could</u> you say goodbye to yesterday?

- D If today was your last day and tomorrow was too late, would you say goodbye to yesterday?
- 6.
- A Against the grain <u>must be</u> a way of life.
- B Against the grain <u>should be</u> a way of life.
- C Against the grain <u>is</u> a way of life.
- D Against the grain <u>ought to be</u> a way of life.
- 7.
- A What's worth the prize is worth the fight.
- B What's worth the prize is <u>always</u> worth the fight.
- C What's worth the prize is <u>surely</u> worth the fight.
- D What's worth the prize is <u>absolutely</u> worth the fight.

8.

- A Every second counts 'cause you have no second chance.
- B Every second counts 'cause you have no second try.
- C Every second counts 'cause <u>there is</u> no second <u>chance</u>.
- D Every second counts 'cause <u>there is</u> no second <u>try</u>.

- 9.
- A So live <u>life</u> like you'll never live <u>it</u> twice.
- B So live <u>life</u> like you'll never live twice.
- C So live <u>it</u> like you'll never live twice.
- D So live \underline{it} like you'll never live \underline{it} twice.
- 10.
- A Don't take the free ride in your life.
- B Don't get a free ride in your life.
- C Don't <u>take the</u> free ride in your <u>own</u> life.
- D Don't get a free ride in your own life.

11.

- A Donate <u>all the dimes</u> you have?
- B Donate <u>each dime</u> you have?
- C Donate every single dime you have?
- D Donate <u>every dime</u> you have?

12.

- A Would you call <u>old</u> friends you never see, reminisce on memories?
- B Would you call friends you never see, reminisce on memories?
- C Would you call <u>your</u> friends you never see and remember memories?
- D Would you call your old friends you never see, and remember memories?
- 13.
- A And would you find <u>the</u> one you're dreaming <u>about</u>?
- B And would you find <u>that</u> one you're dreaming <u>about</u>?
- C And would you find <u>that</u> one you're dreaming <u>of</u>?
- D And would you find <u>the</u> one you're dreaming <u>of</u>?

14.

 A
 Swear up and down to God above that you'd really fall in love?

B Swear up and down to God above that you'd <u>finally</u> fall in love?

C Swear up and down to God above that you'd <u>truly</u> fall in love?

D Swear up and down to God above that you'd <u>quickly</u> fall in love?

15.

A Would you meet the mark by mending a broken heart?

B Would you hit the mark by mending a broken heart?

C Would you make your mark by mending a broken heart?

D Would you leave your mark by mending a broken heart?

16.

A You know, it's never too late to <u>reach for</u> the stars regardless of who you are.

B You know, it's never too late to shoot for the stars regardless of who you are.

C You know, it's never too late to <u>aim for</u> the stars regardless of who you are.

D You know, it's never too late to strive for the stars regardless of who you are.

17.

A So do whatever it takes 'cause you can't rewind <u>a second</u> in this life.

B So do whatever it takes 'cause you can't rewind <u>a minute</u> in this life.

C So do whatever it takes 'cause you can't rewind <u>an instant</u> in this life.

D So do whatever it takes 'cause you can't rewind <u>a moment</u> in this life.

18.

A Let nothing <u>stand</u> in your way 'cause the hands of time are never <u>on</u> your side.

B Let nothing <u>get</u> in your way 'cause the hands of time are never <u>on</u> your side.

C Let nothing <u>stand</u> in your way 'cause the hands of time are never <u>by</u> your side.

D Let nothing <u>get</u> in your way 'cause the hands of time are never <u>by</u> your side.
