DIACHRONIC ANALYSIS OF VERBS AND TENSES IN ENGLISH SONG LYRICS ACROSS GENRES AND TIME (1960-2009): A CORPUS-DRIVEN STUDY

FLORA ANAK GOYAK

SULTAN IDRIS EDUCATION UNIVERSITY

2021

DIACHRONIC ANALYSIS OF VERBS AND TENSES IN ENGLISH SONG LYRICS
ACROSS GENRES AND TIME (1960-2009):
A CORPUS-DRIVEN STUDY

FLORA ANAK GOYAK

DISSERTATION SUBMITTED IN FULFILLMENT OF THE REQUIREMENT
FOR THE DEGREE OF MASTER OF EDUCATION
(RESEARCH MODE)

FACULTY OF LANGUAGES AND COMMUNICATION
SULTAN IDRIS EDUCATION UNIVERSITY

2021

UNIVERSITI PENDIDIKAN SULTAN IDRIS
SULTAN IDRIS EDUCATION UNIVERSITY

Please tick (√)
Project Paper
Masters by Research
Master by Mixed Mode
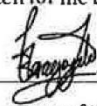PhD

	√

INSTITUTE OF GRADUATE STUDIES

DECLARATION OF ORIGINAL WORK

This declaration is made on the ......16......day of.....February..20...21...

i.   **Student's Declaration:**

I,   Flora Anak Goyak, M20171000334, FBK   (PLEASE INDICATE STUDENT'S NAME, MATRIC NO. AND FACULTY) hereby declare that the work entitled   Diachronic Analysis of Verbs and Tenses in English Song Lyrics across Genres and Time (1960-2009): A Corpus-driven Analysis   is   my original work. I have not copied from any other students' work or from any other sources except where due   reference or acknowledgement is made explicitly in the text, nor has any part been written for me by another person.

_____
Signature of the student

ii.   **Supervisor's Declaration:**

I Assoc. Prof. Dr. Mazura Mastura Muhammad (SUPERVISOR'S NAME) hereby certifies that the work entitled   Diachronic Analysis of Verbs and Tenses in English Song Lyrics   across Genres and Time (1960-2009): A Corpus-driven Analysis

_____(TITLE) was prepared by the above named student, and was submitted to the Institute of Graduate Studies as a * partial/full fulfillment for the conferment of _____Master of Education in TESL_____ (PLEASE INDICATE THE DEGREE), and the aforementioned work, to the best of my knowledge, is the said student's work.

10/3/2021
Date

_____
Signature of the Supervisor

.OF. MADYA DR. MAZURA @ MASTURA BINTI MUHAMMAD
DEKAN
FAKULTI BAHASA DAN KOMUNIKASI
UNIVERSITI PENDIDIKAN SULTAN IDRIS
35900 TANJONG MALIM, PERAK

UNIVERSITI PENDIDIKAN SULTAN IDRIS
SULTAN IDRIS EDUCATION UNIVERSITY

INSTITUT PENGAJIAN SISWAZAH /
INSTITUTE OF GRADUATE STUDIES

BORANG PENGESAHAN PENYERAHAN TESIS/DISERTASI/LAPORAN KERTAS PROJEK
DECLARATION OF THESIS/DISSERTATION/PROJECT PAPER FORM

Tajuk / *Title*:   Diachronic Analysis of Verbs and Tenses in English Song Lyrics
across Genres and Time (1960-2009): A Corpus-Driven Analysis

No. Matrik /*Matric's No.*:   M20171000334

Saya / *I* :   Flora Anak Goyak

(Nama pelajar / *Student's Name*)

mengaku membenarkan Tesis/Disertasi/Laporan Kertas Projek (Kedoktoran/Sarjana)* ini disimpan di Universiti Pendidikan Sultan Idris (Perpustakaan Tuanku Bainun) dengan syarat-syarat kegunaan seperti berikut:-
acknowledged that Universiti Pendidikan Sultan Idris (Tuanku Bainun Library) reserves the right as follows:-

1. Tesis/Disertasi/Laporan Kertas Projek ini adalah hak milik UPSI.
    The thesis is the property of Universiti Pendidikan Sultan Idris

2. Perpustakaan Tuanku Bainun dibenarkan membuat salinan untuk tujuan rujukan dan penyelidikan.
    Tuanku Bainun Library has the right to make copies for the purpose of reference and research.

3. Perpustakaan dibenarkan membuat salinan Tesis/Disertasi ini sebagai bahan pertukaran antara Institusi Pengajian Tinggi.
    The Library has the right to make copies of the thesis for academic exchange.

4. Sila tandakan (√) bagi pilihan kategori di bawah / *Please tick (√) for category below:-*

☐	**SULIT**/*CONFIDENTIAL*	Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub dalam Akta Rahsia Rasmi 1972. / *Contains confidential information under the Official Secret Act 1972*
☐	**TERHAD**/*RESTRICTED*	Mengandungi maklumat terhad yang telah ditentukan oleh organisasi/badan di mana penyelidikan ini dijalankan. / *Contains restircted information as specified by the organization where research was done.*
✓	**TIDAK TERHAD / OPEN ACCESS**	

(Tandatangan Pelajar/ Signature)

(Tandatangan Penyelia / *Signature of Supervisor*)
& (Nama & Cop Rasmi / *Name & Official Stamp*)

PROF. MADYA DR. MAZURA @ MASTURA BINTI MUHAMMAD
DEKAN
FAKULTI BAHASA DAN KOMUNIKASI
UNIVERSITI PENDIDIKAN SULTAN IDRIS
35900 TANJONG MALIM, PERAK

Tarikh: 2/3/2021

Catatan: Jika Tesis/Disertasi ini **SULIT @ TERHAD**, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali sebab dan tempoh laporan ini perlu dikelaskan sebagai **SULIT** dan **TERHAD**.

Notes: If the thesis is CONFIDENTAL or RESTRICTED, please attach with the letter from the organization with period and reasons for confidentiality or restriction.

ACKNOWLEDGEMENT

ABSTRACT

This diachronic study aims to explore linguistics phenomena of verbs and tenses in English song lyrics across genres and time via corpus-driven approach. This study has three objectives; to identify the most frequent lexises, to analyze the use of most frequent verbs, and to analyze the use of most frequent tenses in English song lyrics. A specialized corpus named Diachronic Corpus of English Song Lyrics (DCOESL) consisting of Country, Pop, Rock, and R&B genres from the years 1960 to 2009, was built for the analysis. Linguistics analysis of English song lyrics corpus was used as the research design. Computational corpus instruments were used to generate data. AntConc generated frequency distributions of frequent lexises and adjacent collocations of frequent verbs. CLAWS POS Tagger annotated verbs and tenses. USAS identified semantic categories of verbs. Frequency distributions of verbs and collocations were compared to reference corpora: BNC and COCA. Findings show that DCOESL is saturated with mental verbs (*know*=3562, *want*=2372, *love*=2206) in the present tense, with the highest occurrences in R&B (*know*=1295, *want*=966, *love*=848). Mental verbs occur non randomly across four genres, mean X^2(6, N=8140)=256.5, p<.01. Mental verbs in DCOESL have significant lexical association with mental verbs in the spoken register of reference corpora, mean $G^2_{BNCSPOKEN}$=1047.01, $G^2_{COCASPOKEN}$=2245.31, df=1, p<.01. Mental verbs have strong collocational strength with personal pronouns (*I*=59130, *you=52808*) across genres, mean MI=3.009. Concordances show that mental verbs and personal pronouns co-occur in simple sentence structure to convey perception (*know*), desire (*want*), and affection (*love*), and in the present tense to foster immediacy of communication between an addresser (*I*) and an addressee (*you*). In conclusion, the diachronic analysis of verbs and tenses of English song lyrics has found that mental verbs, present tense, and personal pronouns which are formed in simple sentence structure of English song lyrics, function to express mental activities similar to spoken English. The implication of this study is English song lyrics, especially from the R&B genre, are a potentially authentic teaching resource for teaching spoken English.

ANALISIS DIAKRONIK KATA KERJA DAN KALA DALAM LIRIK LAGU BAHASA

INGGERIS MERENTASI GENRE DAN MASA (1960-2009):

KAJIAN BERDASARKAN KORPUS

ABSTRAK

Kajian diakronik ini bertujuan meneroka fenomena linguistik kata kerja dan kala dalam lirik lagu Inggeris daripada genre muzik dan tahun yang berbeza. Kajian ini mempunyai tiga objektif; mengenal pasti perbendaharaan kata kerap, menganalisis kegunaan kata kerja kerap, dan menganalisis kegunaan kala yang kerap digunakan dalam lirik lagu Inggeris. Korpus khusus iaitu Diachronic Corpus of English Song Lyrics (DCOESL) yang merangkumi genre Country, Pop, R&B, dan Rock, dari tahun 1960 hingga 2009 telah dibina untuk analisis. Analisis linguistik lirik lagu Inggeris telah digunakan sebagai reka bentuk kajian. Instrumen komputasi linguistik digunakan untuk menghasilkan data. AntConc menghasilkan taburan frekuensi leksis kerap dan kolokasi bersebelahan kata kerja. CLAWS POS Tagger menganotasi kata kerja dan kala. USAS mengenal pasti kategori semantik kata kerja. Taburan kata kerja dan kolokasi dibandingkan dengan korpora rujukan umum: BNC dan COCA. Hasil kajian menunjukkan bahawa DCOESL padat dengan kata kerja mental (*know*=3562, *want*=2372, *love*=2206) dalam bentuk kala kini, dengan kekerapan paling tinggi dalam R&B (*know*=1295, *want*=966, *love*=848). Kata kerja mental berlaku secara tidak rawak merentasi empat genre, $X^2(6,N=8140)=256.5$, $p<.01$. Kata kerja mental dalam DCOESL mempunyai kaitan yang signifikan dengan kata kerja mental dalam kategori lisan korpora rujukan, purata $G^2_{BNCLISAN}=1047.01$ and $G^2_{COCALISAN}=2245.31$, df=1, $p<.01$. Kata kerja mental mempunyai hubung kait kolokasi bersebelahan yang kuat dengan kata ganti nama diri (*I*=59130, *you=52808*) merentasi empat genre, purata MI=3.009. Konkordans menunjukkan bahawa gabungan kata kerja dan kata ganti nama diri dalam struktur ayat mudah digunakan untuk menyampaikan persepi (*know*), keinginan (*want*), dan kasih sayang (*love*), dan dalam bentuk kala kini untuk komunikasi secara segera antara pengalamat (*I*) dan penerima (*you*). Kesimpulan kajian analisis diakronik kata kerja dan kala mendapati kata kerja mental, kala kini, dan kata ganti nama diri yang tinggi dalam struktur ayat mudah dalam lirik lagu Inggeris berfungsi menyampaikan aktiviti-aktiviti mental serupa dengan lisan Bahasa Inggeris. Implikasi kajian adalah lirik lagu Inggeris terutama dari genre R&B merupakan sumber yang berpotensi untuk dijadikan sebagai bahan pengajaran lisan Bahasa Inggeris.

CONTENT

APPENDIX A

APPENDIX B

APPENDIX C

APPENDIX D

APPENDIX E

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BNC	British National Corpus
BNC$_{SPOKEN}$	Spoken Register of the British National Corpus
BNC$_{WRITTEN}$	Written Register of the British National Corpus
CEFR	Common European Framework of Reference
COCA	Corpus of Contemporary American English
COCA$_{SPOKEN}$	Spoken Register of Corpus of Contemporary American English
COCA$_{WRITTEN}$	Written Register of Corpus of Contemporary American English
DCOESL	Diachronic Corpus of English Song Lyrics
ESL	English as a Second Language
EFL	English as a Foreign Language
L2	Second Language
LGSWE	Longman Grammar of Spoken and Written English
MI	Mutual Information
NF	Normalized Frequency
PMW	Per Million Words
PTW	Per Ten Thousand Words
RF	Raw Frequency
R&B	Rhythm and Blues
T	T-score

LIST OF SYMBOLS

%	Percentage of Distribution
G2	Log-likelihood
VV0	Base Form of Lexical Verb
VVD	Past Tense of Lexical Verb
VVG	-ing Participle of Lexical Verb
VVN	Past Participle of Lexical Verb
VVZ	-s Form of Lexical Verb
X2	Chi-square

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1   General Introduction

This chapter constitutes descriptions of an overview of songs in the teaching of English language, statement of problem, research objectives, research questions that this study aspires to answer, significance and limitations of the study. This thesis consists of five chapters and concludes with a summary of in each chapter.

1.2   General Overview of Songs Use in Teaching English Language

A song is vocally produced, linguistically meaningful and has melody (Griffee, 1992) and it shares a few common characteristics with speech and poetry. A song also possesses its own unique qualities. Griffee explained that songs have two main features; first, songs convey a lower amount of information than poetry and second,

contain more redundancy than poetry. Griffee (1992) explained that despite they can be heard or they are usually sung, they allow shorter and lighter information. Songs contain more redundancy than poetry (Griffee, 1992) and this redundancy occurs because of the existence of devices such as proverbs, catchphrases and cliché to make songs sound simpler. Songs are relatable to listeners of all age groups and background. The lyrics in songs carry stories from a singer to the audience allowing them to "participate in the songs" (Griffee, 1992, p. 4). These stories are relatable to the audiences from all background.

Schoepp (2001) identifies three main reasons for using songs in English language classrooms namely affective, cognitive and linguistics. Firstly, songs embody an effective teaching tool through which learners can learn language in an entertaining way, particularly acquiring the grammar. Lo and Fai Li (1998) stated that learning English through songs provides a non-threatening atmosphere for learners. Teachers are able to use songs as authentic materials that are intrinsically more active, interesting and stimulating (Shei, 2001, as cited in Rashid and Majid, 2014). The combination of language and entertainment shows to learners that language learning can be fun. Secondly, the repetitive nature of songs promotes memory retention. The rhythm and rhyming patterns in songs are considered one of the effective memory aids as they help students to recall lexis and syntactical chunks (Schoepp, 2001). Thirdly, songs enhance the use of English language that learners will face in real life contexts (Eken, 1996). Incorporating song lyrics in classroom activities provides an opportunity for learners to hear vocabulary and phrases in a natural and meaningful context. The students hear whole sentences and absorb grammar and syntax subconsciously due to the catchy appealing nature of songs (Domoney and Harris, 1993).

Exposure to a language is a vital factor to the success in acquiring a second language (Redwood, 2011, as cited in Fernandez, 2015, p. 10). This is true in the case of grammar especially verb forms whereby students are expected to have a greater mastery over the use of correct tenses in both in verbal and written communications. Hence, it could be argued that exposure to language can tremendously increase the chances for learners to internalize the correct uses of grammar into their daily life. Teachers can increase learners' exposure to the language through the integration of music into the English language classrooms. Additionally, Lull (1992) asserted that popular music has been described as a "form of communication and a sphere of culture that routinely diffuses and amplifies its influence deeply and sensuously into the lives of those who create it, listen to it, use it, dance to it" (Inglis, 1996, as cited in Spicer, 2011, p. 37).

Thus, the current study aims to investigate the various aspects of verbs and tenses used in English song lyrics across time and genres. For this purpose, four music genres; Country, Rhythm and Blues (R&B), Pop and Rock which are selected for this study which are categorized as popular English music genres (Musicmap, 2016), have been selected for this study.

1.3   Problem Statement

To establish effective verbal and written communications, the correct use of verbs and tenses is crucial. Verbs are words that name an action or describe a state of being (Rozakis, 2003). Tense is the manifestation of the time that an action takes place to the moment of speaking (Cowan, 2008). In other words, tense locates an action or

event in a specific moment, while the aspect of a verb is determined by whether the action or event is ongoing or completed.

Non-native speakers have to invest much time to understand and use tenses accurately and appropriately (Sukasame, Kantho and Narrot, 2014). Tense is difficult to master because as generally known, the choice of verb forms for future, past, and present tenses depends on many factors and a rigid set of grammatical rules (Cakir, 2011). For instance, the simple past tense is used in the Question 1 below because the question is about activities, and the work day is considered finished meanwhile present perfect is used in Question 2 because the question is about results and the time at which the question is asked is considered as a continuation of the work day. These tenses create different meanings from the main verb *do* that helps to anchor the listener or reader to understand the meanings behind the sentences.

Question 1: "What did you do at work today?"

Question 2: "What have you done at work today?"

The above examples display how tenses are used to create different meanings from the main verb *do* that helps anchor the listeners or readers understand the meanings behind these questions. It also exhibits the importance of using the correct tenses in ensuring that exact meanings of sentences (or in these cases, questions) are being conveyed.

Comprehending how difficult tenses can be, it is understandable that Malaysian English as Second Language (ESL) learners commit grammatical errors particularly in the uses of verbs and tenses. Based on my experience and observation, the problem in learning verbs and tenses among Malaysian learners is not new. Saadiyah and

Kaladevi (2009) collected a corpus of written essays and discovered that verb tense is one of the most common errors committed by a group of Form 4 students. A total of 346 errors (13.3 per cent) on verb tense and 145 errors (4.7 per cent) on verb form were committed by the 72 participants. The findings show that students committed errors because they encountered difficulty in applying grammar rules. A content analysis by Charanjit, Amreet, Nur and Thilaga (2017) on written essays by a group of Diploma students further proves the difficulties encountered by Malaysian learners. The study shows that tenses as one of the most common committed errors. A total of 226 or 30.4 per cent errors on verb tense were identified in 144 written essays. The two tenses which have not been correctly used and applied in students' writings are present perfect tense (I am wondering how the student in that country were busy to open book) and present continuous tense (That most people that has using) (p. 22). Furthermore, Faridah (1985) employed error analysis to identify the types of verb tense errors made by Malaysian undergraduates of intermediate to upper level proficiency enrolled at the Iowa State University. This study proves that despite the eleven years spent on learning the English language and the score of a minimum 500 on the Test of English as a Foreign Language (TOEFL) before the university admission, these students still encountered problems in applying the correct verbs tenses in their compositions. For instance, 85 and 35 errors were committed on simple present and past tenses respectively. Faridah (1985) stated that some students choose tenses randomly because they do not understand the underlying concepts of these tense (p. 62). These studies have illustrated the evident problems with verbs and tenses faced by Malaysian learners from diverse education stages, proficiency levels, both in urban and rural areas.

Akinwamide (2012) stated that English is a crucial second language (L2) as it is one of the languages that is widely spoken and used in the world including in

Malaysia (as cited in Ismail, Palpanadan and Salam, 2015). In Malaysia, English language is taught as a second language (L2) in all primary and secondary schools. It is recognised as a second language in the Malaysian education planning and policy (Jantmary and Melor, 2014, p. 2). As one of the Asian countries that has been adopting a bilingual education system, the Malaysian government has made it a compulsory subject at all levels of education indicating its position alongside "strong indigenous languages, wide use in speaking, and international outstanding, sometimes official functions, as the language of politics, the media, jurisdiction, higher education, and other such domains" (Jantmary and Melor, 2012, p. 2). Additionally, our current Education Minister has stressed on an education system that focuses on a technology-centric learning with an emphasis on the English language (The Malay Mail, 22 May 2018). Hence, being proficient in the language is important as it is an avenue to acquire knowledge of the world. Hence, a pass in English in Sijil Pelajaran Malaysia (SPM) or the Malaysian Certificate of Education in the national examination taken has been made compulsory for Malaysian students since 2016. This is a disconcerting responsibility for teachers as it was reported in a local newspaper on the study conducted by the Performance Management and Delivery Unit (PEMANDU) of the Prime Minister's Department under the Government Transformation Programme (GTP) in 2013 that:

> a total of 1,191 secondary schools were identified to have SPM English failure rates exceeding 23 per cent…English is students' weakest core subject in national assessments: 25 per cent of candidates failing at the UPSR (Primary School Evaluation Test) level, 23 per cent at PMR level (now called Form Three Assessment - PT3), and 22 per cent at SPM (Malaysian Certificate of Education). The same report also highlighted that Malaysian students perform poorly in international English exams with 50 per cent of candidates failing the 1119. (as cited in New Straits Times, 2015)

The needs for Malaysian learners to be proficient in the language is significant to meet the requirement of the government and schools to acquire knowledge, and

also for future workplace needs (Ministry of Education, 2016). The survey of 3800 respondents consisting of managers conducted by JobStreet (2005, as cited in Lee, 2012), a Malaysian employment agency, revealed poor command of English as the most cited factor of graduate unemployment. Another survey in 2015 involving their clients, managers and senior managers, discovered 64% interviewees failed their job interviews due to similar factor (Employers: Fresh Graduates Have Unrealistic Expectations, 2015). Currently, Malaysian Prime Minister has made it compulsory for top civil servants to undertake English competency tests in ensuring English is practiced and mastered by high level civil servants (The Star. 6 June 2018). These demonstrate that the efforts taken to hone Malaysian learners' English language proficiency.

To achieve Malaysia's educational goals, teachers need to go against all odds to teach effectively. One of the challenges would be for teachers to teach grammar particularly on verbs and tenses. The teaching of grammar is integrated into the four language skills of listening, reading, speaking and writing. With reference to the Ministry of Education (MOE), the Forms 1 to 5 Standard-Based English Language Curriculum (SBELC) for secondary schools provide a guideline on types of verbs and tenses that need to be taught to learners for each school year. For example, Form 1 learners need to master simple present tense and at the end of Form 5, they need to master all the aspects of past, present and future tenses. The grammatical items are infused within the suggested topics for different themes which form the content bases through which the aforementioned grammatical items should be taught.

It is extremely difficult to understand the use of correct verb tenses as it depends on various factors which are not governed solely by grammatical rules (Cakir, 2011). For instances, lexical factors namely frequency and telicity, and phonological

factors such as vowel and consonant, all influence the accuracy of tenses (Horne and Fager, 2015). Hence, teachers need to thoroughly equip themselves and find alternatives to make learning grammar less intimidating to effectively impart knowledge to learners.

The use of authentic materials such as song lyrics is a way to help students learn verbs and tenses. Authentic materials are ordinary texts produced not for language teaching (Carter and Nunan, 2001) but for real-life communicative purposes (Widdowson, 1990). This means authentic materials provide meaningful contexts to bridge the gap between classroom and the real world (Rashid and Majid, 2014, p. 249). Authentic materials assist teachers expose learners to real language, motivate learners, and relate to leaners' needs (Peacock, 1997, as cited in Rashid and Majid, 2014). The current study employed a corpus-driven study of diachronic corpus to reflect the change in language variation and pattern across genres and time especially in the use of verbs and tenses.

The emerging findings on lexical verbs were compared to the British National Corpus (BNC) and Corpus of Contemporary American English (COCA) to investigate the usefulness of the lexical items. The tenses were identified through by keying in specific part of speech (POS) tag labels (see Chapter 3, Table 3.3). A corpus-driven description has "to be comprehensive with respect to corpus evidence" and linguistic categories are derived systematically from the "recurrent patterns and the frequency distributions that emerge from language in context" (Tognini-Bonelli, 2001, p. 87). The discussion was carried out in respect to the emerging findings from the diachronic corpus.

The findings from the current study would be useful for teachers to refer to in using or adapting song lyrics as teaching materials in second language classroom. English song lyrics written by the native speakers provide rich natural occurring language and cultural input. Song lyrics can be used to study verbs and tenses because they provide abundant examples of these grammatical items in genuine contexts. Other than that, songs have been noted as a motivating factor in language classroom (Engh, 2013). This occurs when learners successfully attach meaning to academic lessons (Johnson, 2002). For instance, Travie McCoy's Billionaire lyric relate learners to the context of the song and can be used to write a composition on what they would do if they were a billionaire. The song lyric is also rich in the use of verbs and future tense other than exposing learners to elements of the American culture. Hence, incorporating songs in ESL classrooms would help teachers teach difficult topics (in this case grammar) using authentic materials and for learners, learning grammar will become more meaningful, entertaining and in a natural context.

It is motivating when learning becomes meaningful as students understand the purpose of doing it (Sears, 2003). Indirectly, learning grammar becomes less intimidating because learners are provided with the need of a favorable learning setting of low anxiety, self-confidence and high motivation (Krashen, 1989). Hence, this current study will focus on the diachronic analysis and the use of verbs and tenses in English song lyrics across time and genres and discuss the contribution of English song lyrics to the teaching materials teachers can use in the language classroom to teach verbs and tenses.

Studying the change in variation and patterns of verbs and tenses in the English song lyrics diachronically across genres and time enabled the researcher to identify which of the items are more common throughout the decades and genres for native

speakers. The identified items were attested for statistical reliability. The attested verbs and tenses identified in this study are proven to be vital for native speakers in conveying perception and emotions through song lyrics. Analyses and discussion of findings show that song lyrics can be categorized under spoken English language, particularly conversation which contains addressor-addressee context. Firth (1935) recognized conversation corpora as a vital source of corpus research to really understand "what language really is and how it works" (as cited in Halliday, 2013, p. 24). In Malaysia where the English language is not the first language, English textbooks are provided as a learning source for students. Rühlemann, as cited in Renouf and Kehoe (2009) described that English textbooks generally contain typical formal written registers compared to informal speech like conversation (p. 415).

Hence, corpora that resemble spoken English can be an invaluable source of authentic teaching and learning material for English language teachers and learners. In this study, findings in DCOESL illustrate the features of conversation. To find out the potential of song lyrics as an authentic material for teaching English, a song checklist on attested verbs and tenses from DCOESL was developed and distributed along with a questionnaire via Survey Monkey, an online survey platform. The attested verbs and tenses are found in the list of vital verbs as listed in the Common European Framework of References (CEFR). Feedbacks from teachers through the online survey help to show the potential of utilizing English song lyrics in teaching and learning verbs and tenses. A large number of the respondents utilized the songs checklist to tech speaking skill, particularly conversation (see Section 5.3).

1.4 Research Objectives

This diachronic corpus study aims to investigate the various aspects of verbs and tenses used in English song lyrics across genres and time. This research has three objectives:

1. to identify the most frequent lexises used in English song lyrics across genres and time,

2. to analyze the most frequent verbs used in English song lyric across genres and time, and

3. to analyze the most frequent tenses used in English song lyric across genres and time.

1.5 Research Questions

This study aims to seek answers for the following research questions:

1. What are the most frequent lexises used in English song lyrics across genres and time?

2. How are the most frequent verbs used in English song lyrics across genres and time?

3. How are the most frequent tenses used in English song lyrics across genres and time?

This study utilized mixed-methods approach comprising of qualitative and quantitative methods to answer the research questions. Quantitative method is

followed up by qualitative method, which leads to triangulation of data to establish validity and strengthen confidence in the results (Klette, 2012).

1.6 Operational Definitions

The following subsections provide the important operational definitions used in this study.

1.6.1 Corpus

A corpus is "a body of naturally occurring language consisting of sampled written or spoken texts in machine-readable form" (McEnery, Xiao and Tono, 2006, p. 4). It can be stored as computer folders (Baker, 2010), which allows the use of computers to manipulate and exploit language data (McEnery, Xiao and Tono, 2006, p. 6). The plural of corpus is corpora (Hornby et al., 2010). For example, the British National Corpus (BNC) is a corpus which represents the general British English. Other popular corpora are Corpus of Contemporary American English (COCA), Birmingham Corpus and Lancaster/IBM Spoken English Corpus. This study aims to compile a corpus of English song lyrics across for genres and time called the Diachronic Corpus of English Song Lyrics (DCOESL). DCOESL is made up of 5000 song lyrics of about 1.4 million running words.

Corpora are beneficial tools for language teaching. For teachers, corpora can be employed to create exercises, discuss near-synonyms and collocations as well as to give informed answers to student questions by using examples from 'real language'

(Granath, 2009, p. 49). Aijmer (2009) asserted that corpora can be used to demonstrate variation in grammar. For instance, Roland, Dick and Elman (2007) in their corpus study found that the Subject + Verb + Object (SVO) word order is predominant in the five sets of corpora namely the BNC Written, BNC Spoken, Standard Corpus of Present-Day Edited American English (Brown) Corpus, Switchboard and Wall Street Journal. Other than that, Nishina's (2017) study revealed that sentence patterns of Subject + Verb + Direct Object and Subject + Verb + Direct Object + Indirect Object are the most frequent in a corpus of Pop song lyrics (see Section 2.3), showing that Pop song lyrics are a useful resource as teaching materials to demonstrate variation in grammar.

1.6.2 Corpus Linguistics

Corpus linguistics is a method in studying large collection of naturally occurring texts that are stored as computerized databases. It is a research approach that facilitates empirical investigations of language variation and uses which results greater generalizability and validity in research findings (Biber and Reppen, 2015). Biber and Reppen explained that corpus approach is based on empirical analysis on actual language patterns of large and principled corpora via extensive use of computer programmes. Computer-aided analysis of corpus data makes it possible for researchers to undertake many corpus-based studies involving large computerized corpora. Computer helps to process and manipulate corpus data rapidly at minimal cost, avoids human bias to achieve reliability and allows further automatic processing to be performed for various metadata enrichment (McEnery, Xiao, and Tono, 2006).

The current study compiled and analyzed large song corpus of 5 000 songs from four popular music genres. The lyrics were retrieved from the freely available online database named MetroLyrics, and stored electronically in text (txt.) form for corpus analysis purposes. Data analyses include frequency lists, collocational patterns, concordance lines, and statistical tests of significance. First, frequency lists of verbs and tenses were generated with the aid of AntConc. Most frequent verbs and tenses identified from the generated lists are know, want, and love, and simple present and past tenses respectively. These language items are the major findings in the current study as they show high counts for the two most frequent categories under investigation; verb and tense. Second, adjacent collocations for know, want, and love were searched with the aid of AntConc. Instead of merely focusing on the three aforementioned verbs alone, adjacent collocations assist in giving insights on the uses of the verbs. For instance, the to-infinitive in DCOESL occurs immediately to the right of verb want, which gives characteristic to want from typical lexical and active verb, to semi modal. Real linguistic structure could be identified for the prominent verbs, instead of focusing solely on their 'statistical phenomenon' (Lindquist, 2009, p. 78). Next, concordance lines were generated via AntConc. In this study, concordance lines give the overview on the ways lyrics were composed by native speakers through the use of verbs and tenses. By investigating the sentence structures, common sentence patterns identified from concordance lines contribute to the conclusion that songs use simple sentence structure Subject (S) + Verb (V) + Object (O) in simple present and past tenses to immediately convey perception and emotional connection. Last, online corpus tool, Lancsbox, was utilized to submit know, want, love, verb forms (for tenses), and collocational items to statistical tests of significance. Massive corpus like DCOESL requires computational tool to handle its numerical data, especially ones with complicated formula calculations. The calculations for statistical tests of significance for this study are Mutual Information (MI), t-score, Chi-square (X^2), and Log-likelihood

(G^2). These tests are vital for identifying important verbs and collocations that are not just of high occurrences, but are actually matter linguistically in DCOESL. Manual complex coding for calculations via Microsoft Excel consumed longer time as compared to adopting already coded automatic calculator like Lansbox. The tests results helped the researcher to interpret the test items under investigation (see Subsections 1.6.18-1.6.20).

1.6.3 Corpus-Based and Corpus-Driven

Corpus linguistics can be divided into two approaches; corpus-based and corpus-driven. Corpus-based and corpus-driven approaches are deductive and inductive respectively. Corpus-based linguistics aims to use corpus data to explore a theory or hypothesis in order to "validate it, refute it or refine it" (McEnery and Hardie, 2012, p. 6). This approach derives the validity of linguistic forms and structures from a linguistic theory to be analyzed in the representative corpora. The theory serves as a starting point whereby the findings from the analysis are deduced from the theory itself. The analysis includes the common use of corpus annotation through part-of-speech (POS) tagging or grammatical tagging by assigning POS tag to each word in a corpus (McEnery, Xiao, and Tono, 2006). The results and findings will show the discovery of systematic patterns of use that govern the linguistic features recognized by standard linguistic theory (Biber, as cited in Heine and Narrog, 2010). Substantially, corpus linguistics as a method underpins corpus-based approach as a method.

On the contrary, corpus-driven linguistics utilizes corpus data as an empirical basis from which linguistic constructs emerge from the analysis without any preconceived assumptions. The analysis of linguistic variation from this "radically

different perspective" (Biber, 2009, as cited in Heine and Narrog, 2010) becomes feasible with the accessibility to massive and representative corpora as well as incorporation of computational tools. The prospective corpora are utilized to induce linguistic categories that have not been recognized as the "descriptions aim to be comprehensive with respect to corpus evidence" (Tognini-Bonelli, 2001, p. 84). The linguistic categories are derived systematically from the "recurrent patterns and the frequency distributions that emerge from language in context" (Tognini-Bonelli, 2001, p. 87). Some corpus-driven studies concern the role of frequency evidence especially when it involves the study of lexical bundles while for pattern grammar studies, frequency is not a primary criterion (Heine and Narrog, 2010). Despite the differences in the specific methodological approaches, it can be said that the central aim of corpus-driven approach is to inductively identify new linguistic constructs of a corpus. Hence, the corpus itself encapsulates a theory of language.

The current study employs the corpus-driven approach to analyze verbs and tenses used in English song lyrics across genres and time. Discussion on the emerging results is based on the corpus evidence provided by the current diachronic corpus data.

1.6.4 Diachronic Linguistics

Diachronic linguistics is the study of language change across time (Renouf, 2002). Language change is identifiable and measurable within an existing corpus and changes can operate at any level of textual organization (Renouf, 2002, p. 29). The changes may occur within different aspects of language such as pragmatics, semantics or syntax of the text. For instance, Graham (2016) studied the change in

the use of unparliamentary language of the New Zealand parliament from 1890 to 1950. Graham used Microsoft Access 2010 relational database for corpora analysis and presented the data in the form of frequencies. He discovered that the frequency of unparliamentary terms which reflect lack of respect such as "hypocrite/ hypocrisy/ hypocritical" and "mean/ meanness" experienced dramatic increase by 36 and 27 respectively since 1930 onwards (p. 83). Other than that, Viitasalo (2016) studied the variations for complement patterns of the verb "hope" on written British English. She utilized linguistic annotation and raw frequencies to present data. The findings illustrated that complementation patterns constantly change with some patterns gaining more prominence and while others disappearing (p. 88). For both diachronic studies, researchers are able to trace the development of a language of certain time frame.

The current study is a diachronic study of language to analyze verbs and tenses used in English song lyrics across genres and time. Diachronic study was selected instead of synchronic study, which is the study of language at a particular point of time (Krug and Schlüter, 2013), as the researcher is interested in studying the change in variation and pattern of verbs and tenses in the English song lyrics diachronically across genres and time. For this study, the base form of lexical verbs and simple present tense are constantly highest among their variations. Thus, the lexical verbs and tenses that are of highest occurrences are selected to be investigated in this study.

1.6.5 Content Words

Content words namely adjectives, adverbs, nouns and verbs, are open class words with meaning (Denham and Lobeck, 2009, p. 144). For the current study, the

researcher analyzed the diachronic corpus to discover the frequent content words in the compiled English song lyrics of four music genres across time. The researcher tabulated the frequencies of frequent content words. The main focus of qualitative and quantitative analyzes and discussion is on the base form of lexical verbs (see Section 4.3). It is vital for this study to identify common content words in song lyrics, especially lexical verbs, in order to inform and provide lists of verbs and tenses for English language teachers, and consequently could be adopted or adapted into L2 classrooms for teaching and learning purposes.

1.6.6   Adjectives

Adjectives are words that describe nouns (Denham and Lobeck, 2009, p. 221). In this study, adjectives listed are generated using AntConc for raw frequencies. The discussion on adjectives is not emphasized as this study focuses mainly on verbs and tenses. The C7 tag used for adjectives is JJ.

1.6.7   Adverbs

Adverbs are words that add meanings to verbs, adjectives and other adverbs (Winch, 2005, p. 7) and nouns (Aarts, Chalker and Weiner, 2006, p. 14). This study identified and generated frequency list of adverbs via AntConc. The C7 tag used for adverbs is RR.

1.6.8　Nouns

Nouns are words for a person, place or thing (Denham and Lobeck, 2009, p. 213). In this study, nouns were identified solely to generate their frequency lists via AntConc. The C7 tag used for nouns are NN, NN1, NN2, NNO, NNO2, NNT1, and NNT2.

1.6.9　Verbs

As mentioned earlier, verbs are words that name an action or describe a state of being (Rozakis, 2003). Generally, it is impossible to construct a complete sentence without having a verb in it. Without verbs, sentences will be incomplete because they do not express a complete thought. According to the Cambridge Dictionary of English Grammar, a verb is defined as:

> The PART OF SPEECH which contains words like see, arrive, sleep, discuss, shoot and take off. Verbs are divided into two groups: LEXICAL VERBS (which includes most verbs) and AUXILIARIES (which have some special properties). (Pam, 2013, p. 143)

In this study, lexical verbs were identified and tabulated in frequency lists for analysis. Lexical verbs are any verbs except auxiliary verbs. Auxiliaries or helping verbs be, do, and have will be identified in the texts to help in identifying and analyzing tense forms. Modal auxiliary verbs such as can, could, may, might, must, ought to, shall, should, will, and would were identified in the sample sentences from the texts containing lexical verbs to deduce the meaning of the lyrics.

1.6.10 Lexical Verbs

Lexical verbs comprise of an open class of words that function only as main verbs (Biber, Johansson, Leech, Conrad and Finegan, 1999). Biber et al. (1999) in their study have shown that lexical verbs are the most common class of verb in the Longman Grammar of Spoken and Written English (LSWE) Corpus, with over 100 000 occurrences per million words (pmw) and are extremely common in conversation and fiction (pp. 358-359).

In the current study, the focus is on the base form of lexical verbs. This is because base form of lexical verbs can be inflected to yield four further verb forms namely the third person present, past, -ed participle and -ing participle (Rozakis, 2003). First, the frequent base forms of lexical verbs were identified through the help of computational text analysis. Next, from the tabulated frequencies, the researcher was able to identify the inflection variation of the frequent base forms of the lexical verbs. This enabled the researcher to investigate deeper on the interesting findings emerging from the diachronic corpus instead of focusing on all the forms of verbs due to time constraint.

1.6.11 Diachronic Corpus

Diachronic corpus contains texts from different time periods and have been marked up for year of production (Lindquist, 2009). Diachronic corpus in the current study is intended to reflect the change in verbs and tenses variations across genres and time. For that reason, the researcher assembled song lyrics that expanse for 5 decades in order to analyze a wide range of content words such as adjectives, adverbs, nouns

and lexical verbs in the texts. The researcher focused on analyzing verbs and tenses as these items are the center of this study.

1.6.12 Diachronic Corpus of English Song Lyrics (DCOESL)

The diachronic corpus data in this current study is referred to as DCOESL, acronym for Diachronic Corpus of English Song Lyrics. This corpus consists four music genres namely Country, Pop, Rhythm and Blues (R&B), and Rock. A sum of 5000 songs with every genre contains 1250 songs across 5 decades, from years 1960 until 2009, generating about 1.4 millions words were compiled to build DCOESL.

DCOESL was built with the intention to investigate its linguistics potential as an authentic teaching and learning material for L2 classrooms. The researcher utilized corpus computational tools to generate and identify important language items namely verbs and tenses. The verbs and tenses were submitted for statistical tests of significant namely Chi-square, Mutual Information, and t-score, to ensure that statistical occurrences of the items within DCOESL are not due to randomness. To show the representativeness of the items towards the English language in general, the items were submitted for Log-likelihood tests against general reference corpora namely British National Corpus and Corpus of Contemporary American English. This test helps to identify which English registers do the language items under investigation most likely to resemble. The statistical results from statistical tests of significance aid to interpret the phenomenon of verbs and tenses in DCOESL, and enable discussions on how the two categories can be generalized to features of English language in general to be carried out. For instance, verbs *know*, *want*, and *love* throughout the decades obtained high Log-likelihood scores for spoken English. These items are

listed as important verbs under CEFR to be taught to English language learners. The likeliness to find these items in song lyrics makes it possible to consider songs as an authentic material for teaching and learning for L2 classrooms.

1.6.13 British National Corpus (BNC)

A general corpus built to represent a wide cross-section of British English, both spoken and written, from the later part of the 20th century (Burnard, 2009). In this study, the BNC is used as a benchmark of standard comparison, because it is large enough to represent reliable lexises and grammars items. The frequencies of lexises in DCOESL is compared to the ones in BNC to measure the significance of difference between two scores of the same item. This study uses Log-likelihood (G^2) statistical test of significance to calculate the difference in scores.

1.6.14 Corpus of Contemporary American English (COCA)

The only large and balanced corpus of American English (COCA, 2017). In this study, COCA is utilized as a reference corpus to act as a yardstick of standard comparison of English language used by native speakers. Log-likelihood (G^2) statistical test of significance is adopted to measure the scores of selected items between DCOESL and COCA.

1.6.15 Tenses

Tense is the expression of the time that an action happens in relation to the moment of speaking (Cowan, 2008). In other words, it locates an action or event in a specific moment. The Cambridge Dictionary of English Grammar defines tense as:

> "The GRAMMATICAL CATEGORY which correlates most closely with location in time. Like many other languages, English marks tense in verbs: I see her versus I saw her." (Pam, 2013, p. 134)

In this study, the researcher will identify and analyze the emerging tenses from the diachronic corpus. Below is the table to show the ways tenses are related. Present and present progressive are used to show events that are currently happening. The six past forms show events that took place before the present. The four future forms indicate events that take place in the future.

Table 1.1

English Tenses

Past	Present	Future
Simple past	Simple present	Simple future
Present perfect		Future perfect
Past perfect		
Past progressive	Present	Future progressive
Present perfect progressive		Future perfect progressive
Past perfect progressive		

1.6.16 AntConc

A freeware corpus analysis toolkit for concordance and text analysis developed by Professor Laurence Anthony. This study uses AntConc to generate raw frequency lists, concordance lines, and collocations.

1.6.17 Constituent Likelihood Automatic Word-tagging System (CLAWS)

A part of speech tagger tool developed by UCREL, which is the University Centre for Computer Corpus Research on Language at Lancaster University. In this study, CLAWS C7 tagset is used to identify content words and tenses.

1.6.18 EditPad Pro

A text editor or word processor to clean up texts. In this study, EditPad Pro is used to clean up texts. For example, regular expression \S*_[^VV0]\S* is keyed to identify base form of verbs (VV0), and eliminate other words.

1.6.19 Log-likelihood (G^2) Test

Log-likelihood (G^2) is widely used as a measure of strength of associations (Moore, 2004). This test does not assume that the data is normally distributed. For this study, G^2 functions to provide test scores for lexical items in DCOESL that are selected for comparison with reference corpora namely the BNC and COCA.

1.6.20 Mutual Information (MI) Test

A test that measures collocational strength (McEnery, Xiao and Tono, 2006). In this study, MI scores provide measure to identify significant collocations in DCOESL.

1.6.21 T-Test

A test that measures the significant of a collocate to a node (Hunston,2002). T-score test in this study is used to identify lexical items that are vital for the node words under investigation.

1.7 Significance of Study

This study would substantially benefit teachers in the teaching and learning of grammar particularly the teaching of verbs and tenses. Corpus linguistics is the computer-aided analysis of extensive naturally occurring language data. McEnery, Xiao, and Tono (2006, p. 4) defined corpus linguistics as the study of real life language use in "machine-readable form which may be annotated with various forms of linguistic information". Corpus data have been frequently exploited in the area of grammar studies because they provide a reliable language features analysis (McEnery, Xiao, and Tono, 2006). The use of corpus linguistics in this research benefits teachers who want to understand language use and design effective authentic materials for their students.

The computer aided analysis of corpus data in this research will benefit teachers in many crucial ways including in studying language change over time, language use, designing authentic materials and language testing. Research methodology of this study provides basic computer-based analytic instruments and information on where to find tools namely AntConc, EditPad Pro, CLAWS Part-of-Speech Tagger and UCREL Semantic Analysis System. In a way, this study is "teaching about" and "teaching to exploit" (Leech, 1997, as cited in Wichmann and Fligelstone, 2014) the corpus and basic corpus linguistics skills needed respectively, for teachers who want to equip themselves with knowledge on computer-aided corpus linguistics. The analysis of corpus data in this study could be used by teachers in schools to further equip their knowledge on verbs and tenses. Both quantitative and qualitative corpus linguistic approaches were implemented for the current study. The findings in this research have the potential to expose language teachers on the rich amount of verbs and tenses in song lyrics. The findings were attested for statistical tests of significance and found to be representative of the English language in general, and thus proven to provide reliable qualitative and quantitative data. The qualitative data of the concordance adjunct the quantitative data from the frequency list. Firstly, the main quantitative method used for this study is frequency count. In this study, frequency counts uncovered the frequent lexis used, particularly verbs, which were automatically calculated using AntConc. Secondly, Qualitative method was used to describe variation and pattern of verbs and tenses across genres and time. The analysis of emerging results included comparative analysis of frequency counts of the diachronic corpus with BNC and COCA to investigate the usefulness of the lexical verbs.

The triangulation of these methods will help to provide in-depth understanding on the uses of verbs and tenses in context. Other than that, the diachronic analysis of

verbs and tenses in this study will enable teachers to select and design authentic materials for learners. Teachers need to find ways of exploiting authentic materials in classroom. Authentic materials expose learners to real language and its usage in its natural environment (Kilickaya, 2004). English song lyrics enable teachers to do so by raising learners' awareness of how language works by giving them access to a 'native speaker consultant' in encountering 'real language' rather than made-up examples (Granath, 2009, as cited in Aijmer, 2009). The materials are motivating for learners because authentic texts like English song lyrics are proof that the language is used for real-life communication (Nuttall, 1996). Teachers also could conduct language testing based on the real language data in this study to make data-based judgments about candidates' abilities, knowledge and performance (Alderson, 2000). For instance, teachers could use findings from this study to develop test items to test learners' understanding on the use of verbs and tenses. Teachers have to use more authentic materials in classroom to better prepare learners in using authentic language as textbook alone is inadequate to equip them for real world communications (Otte, 2006, as cited in Rashid and Majid, 2014).

Music is a reflection of the time and place that produced it (Griffee, 1992). Therefore, it will be extremely riveting to explore the ways native speakers convey thoughts and ideas by using verbs and tenses in song lyrics over time. The results from this study will give the researcher deeper insights as a teacher to gain deeper understanding on verbs and tenses used in various contexts. The frequency analysis would help the researcher narrow the focus of grammar item selection of commonly used verb tenses by native speakers that should be taught to learners. Authentic usages of the language from the findings could be effectively used by the researcher to design exercises and demonstrate variations of verbs and tenses to learners.

Besides, the findings will redound to the use of English song lyrics as an authentic material in language classrooms.

1.8   Limitations of Study

There are several limitations in this study namely the lack of general reference of spoken-based corpus data and deficiency of up to date songs lists from recent decade that are more relatable to students. This study selected corpora that consist of spoken English that is not equally proportional to written English for a more detailed comparative analysis. The British National Corpus (BNC) consists chiefly of written register (90 per cent), with spoken register contributes to only 10 per cent of the entire corpus. The Corpus of Contemporary American English (COCA) comprises of mainly written register (80 per cent) and the size of spoken register is 20 per cent of the corpus. This corpus-driven study should consider equally proportioned corpora from the beginning to obtain results that are more generalized to the use of English language. Nevertheless, the findings from this research are supplemented by analysis of concordance lines and statistically significant tests to reach valid and reliable results.

The Diachronic Corpus of English Song Lyrics (DCOESL) consists of four music genres that stretch from years 1960s until 2000s. This study excluded the song samples from 2010s due to the large amount of data that would consume longer time to analyze and discuss. A more up to date data should be included especially for song corpora that have the potential for the purpose of syllabus design and material development. This is because although song lyric remains relatable in term of its context, but music styles change throughout the decades. Students from the present generation might not relate to the same classic Rockabilly music like Elvis Presley's

Jailhouse Rock in 1957 with the current fusion of Rock and Pop from Panic! at the Disco's High Hopes in 2019. Corpora of song lyrics should always be updated with informed data materials that can be utilized as references that are more relatable to the present generation.

1.9   Organization of Chapters

This thesis consists of five chapters. Chapter 1 provides a background on the uses of song lyrics in teaching grammar. This chapter also contains the statement of the problem, research objectives, research questions, operational definitions, significance of the study, limitations, and a summary of the chapter.

Chapter 2 provides detailed review of literature related to the present study. The literature review includes previous studies on the uses of songs in teaching English language, corpus linguistics and types of tenses. At the end of Chapter two, the gaps of the research are presented and the aim of the study are restated.

Chapter 3 explains the methods and procedures that will be employed to collect data and carry data analysis. The chapter will include data description and tools for analysis.

Chapter 4 provides data analysis and discussion. Using mixed-method approach of both quantitative and qualitative methods, the findings will include the answers for the three research questions.

Chapter 5 concludes the entire dissertation. It summarizes the findings from the study and discuss the implications of the study on classroom teaching practice. Recommendation for further studies are suggested before providing a general conclusion.

1.10   Summary

In sum, this chapter informs the benefits of utilizing songs in terms of cognitive and linguistic aspects, for teaching English language based on the perspectives of experts namely Griffee (1992), Schoepp (2001), Eken (1996), and Domoney and Harris (1993). The benefits of integrating songs as an authentic material into second language classrooms could potentially be advantageous for teaching verbs and tenses in Malaysia. Thus, the Diachronic Corpus of English Song Lyrics (DCOESL) was built for the purpose of investigating two language items namely verbs and tenses. Verbs and tenses are reported to be challenging English language aspects for Malaysian learners to master (Saadiyah and Kaladevi, 2009; Charanjit, Amreet, Nur and Thilaga, 2017; Faridah, 1985). In this study, these items were investigated for the ways they are used by native speakers. Based on the findings of this study, vital linguistic phenomenon of verbs and tenses in songs were investigated. The findings were attested for statistical tests of significance to ensure reliable language analysis. List of attested verbs and tenses developed based on DCOESL for the small scale online survey in this study is a step for illustrating a way corpus research aids in developing list of authentic materials, in this case songs containing commonly verbs and tenses used by native speakers to communicate perception and emotions, for teaching L2 classrooms. However, corpus research with metadata like DCOESL needs to be handled with care such as selection of suitable reference corpora and updated corpus data.