

Some morphosyntactic features and variety status of Zimbabwean English: A corpus-based study

Thadeus Marungudzi

Thesis submitted for the degree Doctor of Philosophy in English
at the Vaal Triangle Campus of the North-West University

Promoter: Prof. Bertus Van Rooy

November 2016

DEDICATION

To my mother Jerita Marungudzi nee Chikati

Acknowledgements

Glory to God the Almighty for granting me good health and auspicious days during the course of the research culminating in this thesis.

This thesis would not have seen the light of day without the dedicated and patient guidance and mentorship of my supervisor, Professor Bertus Van Rooy. Thank you for your always timeous, meaningful and enlightening comments at various stages of the writing of the thesis and for always being forthcoming when in various ways, I needed your assistance.

I also want to thank all the individuals and organisations (not mentioned here for purposes of anonymity) who participated in the generation of spoken and written data for this thesis. Your contribution was invaluable.

I would also want to thank Ms. Theresa Louw and Mrs Vongai Madzimore for their assistance with corpus transcription.

To colleagues and friends in the department of English and Media Studies (GZU); Kizito Muchemwa, Golden Maunganidze, Pepukai Chiwewe, Amos Mushati and Tendai Mangena, thank you for your support and encouragement. To my ‘diasporan’ friends; Bob Marimbire, Pepukai Bengura and Mind Dhanere, thank you for believing in me and for cheering me on from across oceans and rivers.

To members in the department of African Languages and Culture (GZU) including Godwin Makaudze, Khensani Madlome, Hellen Shoko and Faith Sibanda as well as Isaac Mhute (ZOU), thank you for the illuminating discussions on aspects of the thesis relating to Zimbabwean indigenous languages and for directing me to appropriate references.

And to my wife, ‘Mai Taa’, thank you for your prayers and for being my unshakeable pillar of support and my never-drying oasis of love in all circumstances. To our three kids, Tariro, Nokutenda and Akudzwe, it’s all because of you. I love you and be blessed.

Abstract

This study aimed to explore the morphosyntactic features characteristic of Zimbabwean English (ZimE) i.e. English as used by Black second language users, and on that basis, determine its variety status. The study mainly adopted the corpus linguistics methodology in which a special corpus of oral and written discourse was collected from various Zimbabwean contexts of use. The corpus was then analysed through *Wordsmith Tools* to determine the extent to which various morphosyntactic features were characteristic of this new English. The morphosyntactic features were in turn explained in terms of possible factors behind their occurrence and compared with features of other peer new English varieties on one hand and Standard English conventions on the other. The study showed that Zimbabwean English in the main shared, to various degrees, a number of features with peer new English varieties in Southern, Eastern and Western Africa. Among the features which were attested in the ZimE corpus were the extension of the progressive aspect to stative verbs, the deletion of *be* before verbs in the progressive form, the deletion of *be* before the auxiliary *gonna*, use of the resumptive pronoun, use of *too*, *too much*, *very much* for *very* qualifier, addition of a *to*-infinitive where Standard English has a bare infinitive, inverted word order in indirect questions, use of *like* as a focussing device. Other features which did not occur in the corpus at all or were extremely rare were also identified. It emerged that, though there are certain features of ZimE which are uniquely Zimbabwean, to a great extent, there were very few areas of morphosyntactic structure in which it departed from Standard English conventions. The attested characteristics of new Englishes or pidgin varieties reported in the World Englishes literature were found to be largely rare or non-existent in the variety altogether. The occurrence of the features attested in the ZimE corpus were explained through factors related to the educational acquisition context of the variety, pragmatic aspects of information-processing as well as cross-linguistic and intralingual factors related to the languages in contact themselves. The variety status question was settled through appeal to the Dynamic Model of the development of new Englishes (Schneider, 2003; 2007) as well as a consideration of the characteristic linguistic features of ZimE attested in the corpus. It emerged that though there is evidence of nativisation at the linguistic level, the absence of codification and acceptance (institutionalisation) of the Zimbabwean variety of English, despite an officially pronounced national indigenisation stance in the socio-economic and cultural life of the country,

are holding back ZimE from progressing beyond the nativisation stage, with chances that the status of the variety is unlikely to change in the foreseeable future.

Key words: morphosyntactic, Linguistic features, Zimbabwean English, new English, interlanguage, language variety, variety status, ICE.

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION	1
1.1 Introduction	1
1.2 Problem statement and contextualisation	2
1.3 Theoretical background	5
1.4 Research questions	9
1.5 Research objectives	9
1.6 Research methodology	9
1.6.1 Broad approach	9
1.6.2 Data collection and analysis	11
1.6.3 Features explored in the study	11
1.7 Ethical considerations	12
1.8 Justification of the study	12
1.9 Structure of the thesis	13
CHAPTER 2 REVIEW OF LITERATURE	14
2.1 Introduction	14
2.2 English as a world language	16
2.2.1 Reasons for the emergence of English as a world language	18
2.2.2 Attitudes towards English as a world language	19
2.3 Study of world Englishes: Theoretical conceptualisations	22
2.4 Research traditions and orientations in the study of world Englishes	27
2.5 The linguistic tradition in the study of world Englishes	31
2.6 Determining statuses of new Englishes	34
2.7 New Englishes and Pidgins and Creoles	37
2.8 African varieties of English: West Africa, East Africa and Southern Africa	40
2.9 The sociolinguistic history and status of English in Zimbabwe	47
2.10 Some studies of English in Zimbabwe	52
2.11 Conclusion	56
CHAPTER 3 RESEARCH METHODOLOGY	58

3.1	Introduction	58
3.2	Historical origins of corpus collection	58
3.3	The influence of technology on corpus linguistics	59
3.4	Motivation for modern corpus collection	60
3.5	Definition of corpora and types of corpora	61
3.5.1	Corpora, archives and other databases	62
3.5.2	Types of corpora	63
3.6	Reasons for building one's own corpus	64
3.7	Basic considerations to be made when building a corpus	65
3.7.1	Recording	67
3.7.2	Transcribing	68
3.7.3	Coding	68
3.8	Limitations in using corpora for research purposes	70
3.9	The possibilities of corpus linguistics in non-native English contexts	71
3.10	Building a corpus to represent a variety of a language	73
3.11	The corpus design of the present study	76
3.11.1	The research questions addressed by the corpus study	76
3.11.2	The corpus sample	76
3.11.3	Data collection procedures	79
3.11.3.1	The collection of spoken data	80
3.11.3.2	The collection of written data	81
3.11.4	Nature of texts and file-naming conventions	84
3.11.5	Corpus analysis techniques	85
3.11.5.1	Frequency lists	85
3.11.5.2	Concordance analysis	86
3.12	Conclusion	88
CHAPTER 4 DATA: PRESENTATION, ANALYSIS AND INTERPRETATION		89
4.2	The eight salient morphosyntactic features of ZimE	90
4.2.1	Extension of progressive aspect to stative verbs (Feature 88)	90
4.2.1.1	Description of the feature in Standard English	90

4.2.1.2	Possible ways in which ZimE and other new varieties of English differ from Standard English	92
4.2.1.3	Results from the corpus	94
4.2.1.4	Interpretation and discussion	96
4.2.1.5	Conclusion	97
4.2.2	Deletion of <i>be</i> auxiliary before progressive (Feature 174)	98
4.2.2.1	Description of feature in Standard English	98
4.2.2.2	Possible ways in which ZimE and other new varieties of English differ from Standard English	99
4.2.2.3	Results from the corpus	99
4.2.2.4	Interpretation and discussion	101
4.2.2.5	Conclusion	102
4.2.3	Deletion of auxiliary <i>be</i> before <i>gonna</i> (Feature 175)	103
4.2.3.1	Description of feature in Standard English	103
4.2.3.2	Possible ways in which ZimE and other new varieties of English differ from Standard English	103
4.2.3.3	Results from the corpus	103
4.2.3.4	Interpretation and discussion	105
4.2.3.5	Conclusion	106
4.2.4	The resumptive pronoun (Feature 194)	107
4.2.4.1	Description of feature in Standard English	107
4.2.4.2	Possible ways in which ZimE and other new varieties of English differ from Standard English	109
4.2.4.3	Results from the corpus	109
4.2.4.3.1	Distance of resumptive pronoun to the head noun	111
4.2.4.3.2	Head noun semantics	113
4.2.4.4	Interpretation and discussion	115
4.2.4.5	Conclusion	116
4.2.5	Addition of ‘ <i>to</i> ’ where StE has bare infinitive (Feature 209)	117
4.2.5.1	Description of the feature in Standard English	117
4.2.5.2	Possible ways in which ZimE and other new Englishes may differ from	

	Standard English	119
4.2.5.3	Results from the corpus analysis	120
4.2.5.4	Interpretation and discussion	122
4.2.5.5	Conclusion	127
4.2.6	<i>Too, too much, very much</i> for ‘very’ as qualifier (Feature 222)	129
4.2.6.1	Description of feature in Standard English	129
4.2.6.2	Possible ways in which ZimE and other new Englishes may differ from Standard English	130
4.2.6.3	Results from the corpus	130
4.2.6.4	Interpretation and discussion	134
4.2.6.5	Conclusion	135
4.2.7	Inverted word order in indirect questions (Feature 227)	135
4.2.7.1	Description of feature in Standard English	135
4.2.7.2	Possible ways in which ZimE and other new Englishes may differ from Standard English	138
4.2.7.3	Results from the corpus	139
4.2.7.4	Interpretation and discussion	142
4.2.7.5	Conclusion	145
4.2.8	Use of <i>like</i> as a focussing device (Feature 234)	146
4.2.8.1	Description of feature in Standard English	146
4.2.8.2	Possible ways in which ZimE and other new Englishes may differ from Standard English	148
4.2.8.3	Results from the corpus	148
4.2.8.4	Interpretation and discussion	152
4.2.8.4.1	Distribution of <i>like</i> as a focussing device across corpus modes	152
4.2.8.4.2	Distribution of <i>like</i> across pragmatic discourse functions	153
4.2.8.4.3	The collocates of <i>like</i> as a focussing device	155
4.2.8.5	Conclusion	155
4.3	Low-frequency/Unattested features	155
4.3.1	<i>She/her</i> used for inanimate referents (Feature 1)	156
4.3.2	<i>Me</i> instead of <i>I</i> in coordinate subjects (Feature 7)	156

4.3.3	<i>Meself/myself</i> instead of <i>I</i> in coordinate subjects (Feature 8)	157
4.3.4	Different count or mass noun distinctions resulting in use of plural for StE singular (Feature 55)	158
4.3.5	Double determiners e.g. demonstrative/article + possessive, with possessive pronoun preposed or postposed (Feature 59)	158
4.3.6	Extension of progressive to habitual contexts (Feature 89)	159
4.3.7	<i>Come</i> -based future ingressive markers (Feature 116)	160
4.3.8	Use of <i>like</i> as a quotative device (Feature 235)	160
4.3.9	Conclusion	161
4.4	Variety status of Zimbabwean English	162
4.5	Conclusion	165
CHAPTER 5 CONCLUSION		166
5.1	Introduction	166
5.2	Summary of findings	166
5.2.1	The morphosyntactic features of ZimE	167
5.2.2	Explanations for the features evident in ZimE	169
5.2.3	The status of ZimE	171
5.3	Contributions of the study	171
5.4	Limitations of the study and suggestions for further research	172
5.5	Overall conclusion	173
References		174
Appendices		191
Appendix 1: Sample of research participant consent form		191
Appendix 2: Information sheet for research participants		193

LIST OF TABLES

Table 2.1:	Research traditions and orientations on the study of world Englishes	28
Table 3.1:	The corpus sample	77
Table 4.1:	The eight salient features of ZimE	89
Table 4.2:	Prevalence of the extension of the progressive to stative verbs	95
Table 4.3:	Distribution of the extension of the progressive aspect across registers	96
Table 4.4:	Prevalence of deletion of the auxiliary <i>be</i> in the progressive in ZimE	100
Table 4.5:	Prevalence of the deletion of auxiliary <i>be</i> before <i>gonna</i>	104
Table 4.6:	Distribution of resumptive and anaphoric pronoun <i>they</i>	110
Table 4.7:	The immediate collocates of the resumptive pronoun in ZimE	111
Table 4.8:	The collocates of resumptive pronouns in terms of head noun semantics	114
Table 4.9:	Prevalence of the bare infinitive in ZimE	121
Table 4.10:	Prevalence of the bare infinitive with other words known to take bare infinitives in Standard English	122
Table 4.11:	Prevalence of the <i>to</i> -infinitive that conforms to Standard use in ZimE	122
Table 4.12:	Prevalence of <i>too</i> , <i>too much</i> or <i>very much</i> for <i>very</i> in ZimE	131
Table 4.13:	Distribution of <i>too</i> , <i>too much</i> and <i>very much</i> as a ‘very’ qualifier across spoken and written modes	132
Table 4.14	Distribution of <i>too</i> , <i>too much</i> and <i>very much</i> as a ‘very’ qualifier across spoken mode registers	133
Table 4.15:	Distribution of <i>too</i> , <i>too much</i> and <i>very much</i> as a ‘very’ qualifier across written mode registers	133
Table 4.16:	Concordance results on indirect questions with inverted word order	141
Table 4.17:	Comparison of word order inversion in indirect questions in spoken and written texts	142
Table 4.18:	Example sentences of <i>like</i> as focussing device (Kortmann & Lunkenheimer, 2012)	147
Table 4.19:	The prevalence of <i>like</i> as a focussing device in ZimE	150
Table 4.20:	Distribution of <i>like</i> as a focussing device in spoken and written corpora	150
Table 4.21:	Pragmatic discourse functions of <i>like</i>	150

Chapter 1

Introduction

1.1 Introduction

This study focusses on the morphosyntactic features and variety status of Zimbabwean English (ZimE). The study comes against the background of a dearth of research on the phenomenon of English in Zimbabwe, in spite of numerous research studies on other new Englishes across the world having been done and on-going. Arua (1998:139) observes that the scantiness of research on new Englishes in Swaziland in particular and southern Africa in general (except South Africa) could be a result of a general assumption that studies of South African English also cover other Englishes in southern African countries. Of course there are indeed commonalities in the substrate languages and contexts of the southern African countries, but there are also differences whose impact on the nature of the variety in each particular case certainly merits research attention. The present study is also motivated by the fact that the few research studies that have been carried out on English in Zimbabwe have largely focussed on the functional status of English from a language planning and policy perspective e.g. Fitzmaurice (2010), Hungwe (2007), Kadenge (2010), Kadenge and Mugari (2015) Kadenge and Nkomo (2011), Magura (1984), McGinley (1987) , Magwa (2010a), Magwa (2010b), Makoni (1993), Marungudzi (2016a), Mlambo (2009) and Ngara (1977; 1982). The studies just mentioned give little or no attention to the internal linguistic structure of the variety, an aspect which I believe, as I point out in Marungudzi (2016b:7), is paramount in the determination of the ontological status of ZimE. In this study, and particularly focussing at the morphosyntactic level the linguistic structure of the Zimbabwean variety of English is explored through corpus linguistic methods, where oral and written texts from different spheres of use were collected and analysed using Wordsmith Tools.

This chapter introduces the study by giving necessary background information, explaining the research problem, outlining the research questions, stating the research objectives, giving a preliminary review of related literature, summarising the research methodology as well as outlining the structure of the rest of the research report. Below I explain the research problem.

1.2 Problem statement and contextualisation

The relocation from Europe and subsequent settlement of English-speaking populations across the globe witnessed during the colonial period has led to the development of different varieties of English on a global scale (Schneider, 2003:235). A number of terms have been used to refer to these new varieties of English, among them world Englishes, foreign English varieties (Bolton, 2005; Van Rooy, 2011), global Englishes or postcolonial Englishes (Schneider, 2003; 2007). As studies, by Gisborne (2000), Gut (2007), Kruger and Van Rooy (2016), Mesthrie (1997; 2004; 2006), Schmied (1991; 1996), Van Rooy (2011; 2014), Van Rooy and Terblanche (2006) and Van Rooy *et al.* (2010) among many others, illustrate, these new Englishes show both structural and functional differences between themselves on the one hand; and between themselves and native varieties of English on the other. Gut (2011) notes that these differences are a result of contact between English and other languages. It is this contact that results in the indigenisation and nativisation of the English language. According to Schneider (2003:233), English is a language that is currently “growing roots in a great many countries and communities around the world, being appropriated by local speakers, and in that process it is diversifying and developing new dialects...”. As Schneider (2011:3) points out, the process of indigenisation and nativisation of English “is a product of the very recent past and not primarily of the colonial heritage of former colonies in the British Empire”, a rather ironic state of affairs.

The status and acceptability of these emerging Englishes as institutionalised varieties have, however, remained controversial. There are varieties that are deemed to have actually achieved conventional variety status e.g. Singapore English, or that have at least reached the nativisation stage (stage 3) of Schneider’s dynamic model e.g. Indian English, Hong Kong English and Malaysian English, though the latter three appear to have ‘fossilised’ (Schneider, 2003:250-252). However, there are a number of varieties whose features and statuses remain indeterminate. This is partly a result of the fact that the sociolinguistic milieu and the linguistic features of such new Englishes have not yet been fully described (Marungudzi, 2016a:9). There are cases where studies of these new varieties have been done, though this has been done in a piecemeal way with the aim being usually to compare the varieties to some Standard English (Van Rooy, 2008). From a methodological point of view, many of these studies relied on anecdotal evidence, except for varieties where corpora were available. Furthermore, the question of assigning status to a

particular language variety is admittedly a complicated one, hinged as it is in addition to linguistic factors, on a number of other factors which actually have little or nothing to do with language (Gut, 2011; Schneider, 2003; 2007; 2011; Van Rooy, 2011). Thus this study seeks to explore and explain the nature of the linguistic features of one of the lesser-known varieties of new Englishes and address issues to do with its status. Kortmann and Lunkenheimer (2012) have compiled the linguistic features of some 74 varieties of English distributed across the Anglophone world. There is no reference to the Zimbabwean variety of English as used by black second language users in this project, though work by Fitzmaurice (2010; 2012) focuses on English as used by white Zimbabwean speakers, which is only a historical variety that is no longer spoken to any notable extent in the country.

In his dynamic model on the evolution of new varieties of English, Schneider (2003) postulates that there are five evolution stages through which all speech communities experiencing a new English go, namely the foundation phase, the exonormative stabilisation phase, the nativisation phase, the endonormative stabilisation phase and the diversification phase (also known as the differentiation phase). Each of the five stages of the dynamic model are described in detail in Chapter 2, but suffice it here to say this model wields remarkable explanatory power in regard to the linguistic features and variety status of ZimE, or any other new English for that matter, in the sense that it is possible to place the variety in any of the five phases depending on its structural characteristics as well as its sociolinguistic history and milieu. However, in spite of its successes in explaining Outer Circle varieties, the dynamic model has been found inappropriate in explaining Expanding Circle varieties, resulting in Schneider (2014) proposing the ‘transnational attraction’ framework in such contexts.

The purpose of this study is to explore the linguistic features and variety status of the English used by Black second language users in Zimbabwe. The study comes against the backdrop of little research on ZimE in general as well as the meagre attention given to the linguistic features of the variety, particularly at the morphosyntactic level. Furthermore in the majority of studies on ZimE there has been reliance on anecdotal evidence, conjecture and intuitive introspection to determine the linguistic features, a point that has already been mentioned and will be elaborated in the following chapter.

The study seeks to give a description of the morphosyntactic features of the English used in Zimbabwe by black second language users, determine the distribution of those features and explain their possible origin. The rationale for the argument and the approach taken in the present study is that the ontological status of ZimE can be settled more reliably through an exploration and explanation of its linguistic features. It is also my strong conviction that the corpus analysis method is so far the most appropriate and most reliable method to ascertain the linguistic features characteristic of ZimE as it goes beyond mere sociolinguistic introspection or questionnaire responses to determine the linguistic characteristics of a language variety, as was the case with previous studies on English in Zimbabwe. Once the key question of the linguistic features of ZimE is addressed, the answer to another critical question of whether ZimE is indeed a new variety of English or a mere interlanguage will be determined. These two questions are addressed through a method where a comprehensive collection of different types of texts of the Zimbabwean variety of English are explored using corpus linguistic methods rather than the traditional and often inaccurate and unreliable methods of using intuition and/or anecdotal evidence employed in previous studies.

According to Reppen (2010:31), “a corpus can serve as a useful tool for discovering many aspects of language use that otherwise may go unnoticed”. In addition, questions related to aspects of how language use varies by situation or over time, are also ideal areas to explore through corpus research (Reppen, 2010:31). The corpus approach has been employed to a large extent, in different research studies on Black South African English (e.g. Botha, 2012; de Klerk , 2003; Kruger & Van Rooy, 2016, Mesthrie, 1997; 2004; Van Rooy, 2009; Van Rooy, 2011; Van Rooy, 2014; Van Rooy & Kruger, 2016; Van Rooy & Terblanche, 2006; Van Rooy *et al.*, 2010; Wade, 1996) yielding illuminating insights into the structure and use of African (especially South African varieties of English). Section 1.7 below elaborates on the corpus-based approach but before we turn our attention to that, it is important to chart the theoretical background to the study of new Englishes. (More on the theoretical background will be discussed in section 2.2 and 2.3 of the following chapter).

1.3 Theoretical background

Scholars have proposed various approaches to the study of new Englishes. Adopting a diachronic approach, Schneider (2003; 2007) views new Englishes as emanating from the appropriation of the English language by different speech communities. As noted in 1.1 above, the emergence of these new Englishes is a product, not only of the colonisation process, but a phenomenon that mushroomed more noticeably in the period after the colonial settlers had actually departed. Schneider (2003:234) goes on to observe that despite the dissimilarities among the resultant varieties, the varieties follow a uniform developmental process shaped by consistent sociolinguistic and language contact conditions. Schneider (2003:234) argues that the structural and sociolinguistic similarities among the new Englishes, despite the different settings, “are products of fundamentally similar contact processes, to be accounted for by theories of communication, accommodation and identity formation”.

Gut (2011) also observes that the features of new Englishes are a result of cross-linguistic influence - a highly complex phenomenon that often does not surface as direct structural transfer (as previously believed) but rather as the use of prior linguistic knowledge that can manifest itself in such diverse ways as the speed and path of acquisition, the avoidance or the overproduction of certain structures. Gut (2011:119) proposes that the process of innovation consists of the following steps, in the sequence given:

- Some features of the indigenous language form part of the language productions of individual speakers who are learning English.
- These features remain in the speakers’ language productions even if they have attained a high level of competence in English.
- These features are adopted by speakers of subsequent generations, some of whom might acquire English as a first language.

In conclusion, Gut (2011:119-120) posits that:

- For each innovative structure in new English varieties, different investigations of the different ways in which it might have emerged are necessary.

- The classification of a linguistic structure as an innovation or error proceeds on extra-linguistic rather than linguistic grounds.
- Errors and innovations should therefore not be categorised by linguists as distinct from each other but rather as structures representing two end-points of a continuum.

This agrees with Van Rooy's (2011) observation that genuine linguistic conventions emerge from forms that may have started as errors, particularly errors in the form of undue analogy and overextension. With time, however, such errors ascend to the status of conventionalised innovations through benign neglect, and subsequent ability of the speakers of the new English to deliver their message among themselves and to the outside world using those innovations (de Klerk, 1999) or through widespread use in a speech community, larger geographical diffusion, codification and acceptance by authorities (Bamgbose, 1998).

On the other hand, Mesthrie (2003) advocates, as a starting point, the study of the establishment of individual world Englishes and observes that regional dialects of settlers, sailors' sociolects, soldiers' languages and the L2 of many missionaries were the grist that fed into the mill that processed and produced new varieties. Mesthrie (2003: 450) proceeds to identify three other approaches that may be adopted in the study of world Englishes. These are:

- establishing a truly comparative data-base for linguistic analysis,
- refining our tools for describing and accounting for variation, and
- describing language shift where it is taking place.

It is clear here that the approach by Schneider (2003; 2007) concurs with that of Mesthrie (2003), particularly in respect of the first recommendation i.e. to start by studying the establishment of individual new Englishes. Schneider's approach is also a departure from that of the static three circles model (Kachru, 1982; 1986).

Taking a cue from Mesthrie (2003), as it were, Van Rooy and Terblanche (2006), Van Rooy (2008) and Xiao (2009) adopt a synchronic approach to the study of new Englishes and invoke the multidimensional analysis model, a model used by Biber (1988) in the study of register variation in English. It can be surmised here that the three are making an attempt at refining our tools for describing and accounting for linguistic variation.

According to Xiao (2009), world Englishes have been studied from two broad perspectives; the sociocultural perspective (to which Kachru and Schneider's approaches belong) which is concerned with the elaboration of theories and models of development, spread, classification and interaction of new Englishes; and the linguistic perspective which focuses on the linguistic features in selected varieties of English at various levels including phonetics and phonology, morphology, syntax and discourse. This view concurs with the approach proposed by Mesthrie (2003) i.e. "establishing a truly comparative data-base for linguistic analysis" as noted above. For the purposes of this study, it is by and large, the linguistic approach that is going to be adopted. The variety of English that is used in Zimbabwe as represented in both oral and written texts produced by Black second language users will be analysed for morphosyntactic features.

Giving what can be regarded as a bird's eye view of the approaches taken in the study of new Englishes, Bolton (2003; 2005) identifies eleven approaches in these studies spanning from the Pidgin and Creole Studies of the 1930s to the Linguistic Futurology of the late 1990s and the present. These comprise the following:

- English studies
- English corpus linguistics
- Sociolinguistic approaches
- A features-based approach
- Kachruvian studies
- Pidgin and Creole studies
- Applied linguistics
- Lexicography
- Popularisers
- Critical linguistics and
- Linguistic futurology

Bolton (2003; 2005) and Schneider (2007) detail the tenets of each of these approaches but an important observation that can be made in regard to the approaches is that there are certainly areas of overlap (Bolton, 2005:69) between them. They note that the first two approaches display a linguistic orientation while remainder adopts a sociolinguistic orientation.

However, despite acknowledging the comprehensiveness of Bolton's synthesis of the approaches, Schneider (2007:11-12) has criticised it on the score that "it compiles subdisciplines which are not really on a par". Commenting on these approaches, Schneider (2007:15) suggests that the approaches can be classified into those that offer attention to linguistic structure ("which focus primarily upon the structural properties of given languages and language varieties on the levels of phonology, lexis and grammar" as distinct "from those which are interested in non-structural correlates and conditions of language use in a society") and those with "individual details and case studies" as their goals as distinct from "those that aim at broader generalisations of some kind".

In the present study, the morphosyntactic features of a specific variety of English, i.e. ZimE as used by Black second language users, as well as its status will be explored through reference to a comprehensive corpus with the aim of advancing a pluricentric approach to world Englishes. Against this background, the approach adopted here may be said to be informed by about four traditions: the English Studies Approach, the English Corpus Linguistics Approach, the Features-Based Approach and the Kachruvian Approach. That this study is informed by a mosaic of perspectives would not be surprising because, as noted above, Bolton (2005:69) offers the caveat that there are significant overlaps among them. More importantly and as noted earlier, an adoption of the linguistic approaches is of central significance to the determination of the ontological status of ZimE.

In terms of the synchronic-diachronic dimension, the approach adopted here, like the approach taken by Buregeya (2006), Makalela (2004), Van Rooy (2008) and Van Rooy *et al.* (2010), can be said to be synchronic i.e. it focuses on an array of the linguistic features of a specific variety at a given time. Diachronic studies in new Englishes research are relatively rare though more recently Collins *et al.* (2014), Evans (2015) Van Rooy and Piotrowska (2015), and Van Rooy and Wasserman (2014) have adopted the approach.

1.4 Research questions

To resolve the research problem outlined in section 1.3 above, the following questions will be posed and addressed:

- What morphosyntactic features are characteristic of Zimbabwean English as used by black second language users?
- How can the morphosyntactic features of Zimbabwean English, as used by black second language users be accounted for?
- What is the status of Zimbabwean English as used by black second language users?

1.5 Research objectives

The research questions will be answered by attempting to meet the following objectives:

- Collect a corpus of Zimbabwean English as used by black second language users, and identify its characteristic morphosyntactic features.
- Account for the characteristic morphosyntactic features of Zimbabwean English as used by black second language users in particular.
- Determine the 'new variety' status of Zimbabwean English as used by black second language users.

1.6 Research methodology

1.6.1 Broad approach

Broadly speaking, the present study largely used the corpus-based approach to address the research questions. Previous research largely relied on isolated observation of examples and introspection. As indicated earlier, the corpus-based approach adopted in this study was deemed most appropriate as it goes beyond introspection and observation of isolated examples used by previous researchers on ZimE or the questionnaire method used in the compilation of *The*

Mouton World Atlas on Variation in English. A corpus-based approach is an approach to the study of language use that has the following strengths:

- it is empirical, and analyses the actual patterns of language use in natural texts;
- it utilises a large and principled collection of natural texts, known as a corpus as the basis for analysis;
- it makes extensive use of computers for analysis, using both automatic and interactive techniques; and
- it depends on both quantitative and qualitative analytical techniques. (Biber *et al.*, 1998:4)

Because it uses computers extensively, the corpus-based approach has the following advantages:

- computers make it possible to identify and analyse complex patterns of language use;
- computers allow the storage and analysis of a larger database of natural language that could hardly be dealt with by hand;
- the corpus-based approach provides consistent, reliable analyses because computers don't change their minds or get tired; and
- an interactive strategy can be adopted, allowing the human analyst to make difficult linguistic judgments while the computer takes care of record-keeping. (Biber *et al.*, 1998:5)

However, the goal of the corpus-based approach is not simply to report quantitative findings, but “to explore the importance of these findings for learning about the patterns of language use” (Biber *et al.*, 1998:5) and such exploration will be undertaken in the present study. A corpus-based approach allows researchers to identify and analyse ‘association patterns’ i.e. the systematic ways in which linguistic features are used in association with other linguistic and non-linguistic features (Biber *et al.*, 1998:5). Schmieid (1990:259) also observes that one advantage of the corpus-based approach is that it allows the researcher “to confirm quantitatively and statistically impressions he has gained from introspection or participant observation about qualitative differences between native and non-native Englishes”.

1.6.2 Data collection and analysis

In order to address the research questions, I collected a balanced corpus of oral and written texts as produced by black second language users of ZimE. The texts collected were selected following, partially though, the International Corpus of English (ICE) model. The ICE is designed for the synchronic study of world Englishes and consists of 60 per cent spoken discourse and 40 per cent written discourse. In the context of the available time and budget, I collected 195 texts in total. Each of the texts was at least 2 000 words long, adding up to a corpus of 390 000 words. The texts came from a wide range of genres. Section 3.11 gives more detail regarding the characteristics of the corpus.

The corpus analysis focussed primarily on establishing the distribution of morphosyntactic features selected from 13 domains i.e. pronouns, noun phrase, tense and aspect, modal verbs, verb morphology, voice, negation, agreement, relativisation, complementation, adverbial subordination, verbs and prepositions, and discourse organisation and word order as proposed by Kortmann and Lunkenheimer (2011; 2012).

1.6.3 Features explored in the study

As indicated in section 4.1, a total of 8 features (2 from the agreement category, 2 from the discourse organisation and word order category and one from each of the verb tense and aspect, complementation, relativisation and adverb and preposition categories) were included for exploration. Thus, the features selected were fairly representative of the full range of morphosyntactic features identified by Kortmann and Lunkenheimer (2012). The 8 features had been found through the concordance function of Wordsmith Tools to be at least neither rare nor pervasive in the ZimE corpus (at least Grade B, in terms of Kortmann & Lunkenheimer's (2012) taxonomy). The rationale for the selection of features that were at least neither rare nor pervasive was that such features would be most characteristic of the Zimbabwean variety of English. To establish the linguistic identity of ZimE a bit more comprehensively, brief comments are also made on six other features that were found through the concordance function to be extremely rare in the corpus. These features constitute linguistic forms that negatively

characterise ZimE. Morphosyntactic features that did not occur in the corpus were not explored in the analysis. Chapter 3 gives more detail on methodological issues.

1.7 Ethical issues

In order to collect all the forms of data indicated above, there was need to seek consent from the data providers, and in order to abide by this ethical requirement, advance permission to collect the data was sought. Since the study entailed recording of spoken language, a recording consent form was designed and given to all research participants to complete before recording sessions. It was also ascertained that the data would be treated confidentially and used only for the purposes of the research study and for future use aligned with the general protocols used for corpora in the public domain, a fact that was made clear to the data providers.

1.8 Justification of the study

This study seeks to contribute to the pluricentric study of new Englishes by exploring the morphosyntactic features and variety status of one of the lesser-known varieties of English. In the context of the absence of data on ZimE in international corpus projects such as the International Corpus of English as well as scanty research on the variety, this study is bound to go some way in closing that gap. The study also proposes to account for the morphosyntactic features of black ZimE and their distribution in different genres. The study finally seeks to determine at which stage of development (of the Dynamic Model of Schneider (2003; 2007)) ZimE, as used by second language users is, and predict the probable future development of the variety. Thus, overall, like any other corpus linguistics study, this study also casts light on processes attendant upon language contact, development and change in general.

1.9 Structure of the thesis

This chapter has contextualised the study by, in the main, outlining the research problem and situating it within the broader context of the study of new Englishes and more narrowly of the study of ZimE. Chapter 2 focusses on a review of related literature both in terms of theoretical framework and research traditions in new Englishes as well as research studies that have been carried out on mainly African Englishes and on ZimE. Parts of the chapter thus engage the question of the variety status of the Zimbabwean variety of English, to which the final part of Chapter 4 returns. Chapter 3 paves the way for Chapter 4 (the findings chapter in which the salient features of ZimE and their explanations are reported) by describing in broad terms the corpus linguistics approach, and specifying the procedures of data collection and data analysis. Chapter 5 concludes the research report by summarising the findings, indicating the limitations, and offering suggestions for further research.

Chapter 2

Review of literature

2.1 Introduction

This chapter reviews the literature related to the study of English as used by second language speakers in Zimbabwe. It gives an overview of the theoretical approaches to the development and study of new Englishes, the relationship between new Englishes and the related phenomenon of pidgins and creoles, the processes leading to the emergence of new Englishes, the general grammatical features of English as well as those of African varieties of English that have been explored to date. The chapter also focuses on the history of the role and status of English in Zimbabwe before previewing some studies that focus on the study of English in Zimbabwe. Thus the chapter seeks to situate the analysis of the morphosyntactic features and variety status of Zimbabwean English within the broad context of the study of world Englishes.

Section 2.2 focuses on the phenomenon of English as a global language, outlining the reasons behind its rise as a global language as well as the attitudes of host countries towards the presence of the English languages in their countries. This is clearly a necessary background as it provides a broader historical context to the object of study, i.e. morphosyntactic features and variety status of ZimE, a consequence of the global spread of English and its subsequent entrenchment in different geopolitical territories across the world.

The study of the morphosyntactic features and variety status of ZimE falls under the general study of World Englishes. In section 2.3, I consider the theoretical approaches and models to the study of World Englishes from the Three Concentric Circles model (Kachru, 1982) to the Dynamic Model (Schneider, 2003; 2007). A discussion of these frameworks and what each of them entails provides a necessary background and justification for the corpus method that is used in the present study in the sense that the corpus method seeks to complement these sociolinguistic models.

Closely related to the theoretical approaches and models are research traditions and orientations which I outline in section 2.4. This section helps trace existing research traditions in the study of

world Englishes and highlight their strengths and shortcomings so that the place and relevance of the linguistic approach adopted in the present study is established.

Section 2.5 focuses on some global projects in the study of world Englishes which have specifically addressed the nature of English from linguistic perspectives, namely the *electronic World Atlas of Variation in English* (eWAVE), the *Atlas of Pidgin and Creole Language Structures* (APiCS) and *The Mouton Atlas of Variation in World English*. A discussion of these projects provides background in terms of the complex of features whose use can be possibly explored, possible methodological strategies in terms of data treatment as well as in terms of findings with which results on ZimE may be compared.

To lay the ground for answering the all-important new English variety status question, section 2.6 discusses theoretical views regarding how variety status may be ascribed. Here the distinction between error and innovation (Van Rooy, 2011) as well as the conceptualisation of criteria for ascribing new English variety status by Bamgbose (1998), Mollin (2007) and Schneider (2003) are discussed. This discussion is important in the sense that the variety status question may be overall resolved through reference to these explanations.

Two other phenomena are also related to new Englishes, namely pidgins and creoles, which are presented in section 2.7. A discussion of the formation processes and characteristics of pidgins and creoles as well as how they relate to new Englishes provides useful background against which particular features of ZimE may be explained. Pidgins and creoles are language varieties born out of language contact processes, just like new Englishes. The discussion lays the basis for drawing possible comparisons between them and new Englishes, engendering a closer understanding of language evolution processes in general and to some extent helping characterise ZimE in particular.

Since the study of new Englishes is not something new in the world, and indeed in Africa, section 2.8 focuses on the general phenomenon of African Englishes as well as more specific studies in which African varieties of English were explored. Through a review of the specific features that the selected studies focus on, a sense of what may be expected from an analysis of ZimE may be established, making it possible to make comparisons between the features of those varieties and those of ZimE.

Section 2.9 focusses on the sociolinguistic status of English in Zimbabwe during the colonial and postcolonial periods. Where applicable, the sociohistorical episodes that coincide with the phases of Schneider's (2003; 2007) dynamic model will be noted. This juxtaposition will help shed light on the attendant linguistic characteristics of each developmental stage and build towards answering the ultimate question of the variety status of ZimE. The section also refers to language policy issues which are partly responsible for the origin and development of localised English varieties. Overall it should be pointed out that an understanding of the sociolinguistic context of ZimE enables us to explain the functional purpose of the linguistic forms that will be identified.

In section 2.10, a case for the importance of exploring the linguistic aspects of ZimE is made through reference to the dearth of research in this vein and a discussion of early studies of ZimE. It is also shown in this section that the determination of the status of a variety by a majority of previous researchers in the context of Zimbabwe English was mainly based on isolated personal observations and introspective conjectures, arguably inadequate means to determine the status of the variety of English. This discussion thus justifies the use of an alternative methodology, the corpus method, as carrying more potential to resolve the identity and variety status question of ZimE.

2.2 English as a world language

It should be pointed out at the outset that the exploration of the linguistic features and variety status of English is indebted, in the main, to the presence of the English language in Zimbabwe. As Phillipson (1992:23) points out, English is no longer of concern only in north-west Europe, America, Australia and New Zealand ('the antipodes') because it is now entrenched worldwide. The presence of English in Zimbabwe and indeed in other parts of the world is a result of its global spread (Crystal, 2006; Phillipson, 1992; Schneider, 2011) "through British colonialism, international interdependence, 'revolutions' in technology, transport, communications and commerce, and because it is the language of the USA, a major economic power..." (Phillipson, 1992:23). Of course, this global spread has had dire consequences on the fortunes of the indigenous languages of the host countries (Phillipson, 1992: 17; Skutnabb-Kangas, 2000; 2001; 2002; Kontra *et al.*, 2001). According to Skutnabb-Kangas (2002:4), "even if all official

languages are threats to non-official languages, English is today the world's most important killer language". This section thus offers a necessary background to a discussion of the linguistic features and variety status of English as used by second language learners in Zimbabwe. According to de Klerk (1999:311), a study of the historical roots for new Englishes is necessary because all new Englishes need to be defined ontologically rather than phylogenetically. Furthermore, as further argued in 2.9 below, the linguistic features and variety status of the English used by second language users in Zimbabwe owe their existence to the special role that the English language plays in Zimbabwe. As will be seen Chapter 4, the question of variety status in particular is related to individual and community attitudes towards the English language which are outlined in 2.2.2 below.

Specifically, in spite of promulgations in the 20th amendment of *The Constitution of Zimbabwe* that some sixteen languages spoken in Zimbabwe are now 'officially recognised', the English language continues to dominate official government administration, the media, parliamentary deliberations, education, and science and technology. It is this widespread use, specifically in formal settings, that provokes the question as to whether the English has not now been 'tamed' as it were, to reflect structural and cultural aspects of the indigenous languages of Zimbabwe. According to Phillipson (1992:24), whether the English of core-speaking countries and that of periphery countries form one language or several is an essential question given that each variant functions in its own multilingual ecosystem. This section focuses firstly on the reasons behind the global spread of the English language and then moves to how recipient countries of the sojourning English language have reacted to its presence in the countries.

Crystal (2006:423) states that a language achieves a truly global status when it develops a special role that is recognised in every country. Against this background, Crystal goes on to argue that English is indeed a world language in the sense that, in addition to it being spoken by large numbers of people as a first language as happens in the USA, Canada, Britain, Ireland, Australia, New Zealand, South Africa and a number of territories in different parts of the world, it also has some kind of special administrative status in over 70 countries that include Ghana, India, Nigeria, Singapore, Vanuatu and Zimbabwe. English also has a special role in over 100 countries in Africa, Asia, Europe and South America where it has been made a priority in the countries' foreign-language teaching policies though it has not been accorded official status.

According to Crystal (2006:422), though there is a general understanding of the factors governing the presence and fate of a global language (such as is presently the case with English or was previously the case with Latin or is likely to become the case with Spanish, Chinese, Arabic and Hindi), “we have very little understanding of how they interact, and of what happens to the structural character of a language when it achieves a global presence”. The investigation of English as a world language therefore “provides a fresh testing ground for sociolinguistic hypotheses which previously had only regional validity and a domain where we may encounter new kinds of phenomena which might one day motivate a global reconceptualisation of that subject” (Crystal, 2006:422).

2.2.1 Reasons for the emergence of English as a world language

Popular, if uninformed opinion often (see discussion of Schmieid, 1985 below) takes the global spread of English as traceable to its intrinsic linguistic structure such as its ‘simplicity’; as evidenced by ‘a relative lack of inflectional endings, absence of grammatical gender and lexical tone, non-use of honorifics etc’, arguing that such factors make it attractive or easy to learn (Crystal, 2006:426), confirming the observation by Schmieid (1991:164-165) that “expressions of negative or positive feelings towards a language may reflect impressions of linguistic difficulty or simplicity, ease or difficulty of learning, the degree of importance or status it has in the community or even the importance of people who use it as a first or second language”. Crystal observes that such accounts ignore the syntactic, lexical and stylistic complexities that characterise English. Furthermore, linguists are generally agreed that all languages are equal and are capable of expressing any idea no matter how complex (Akmajian *et al.*, 2010: 9-10). In fact “a language’s global stature has nothing to do with linguistic character” (Crystal, 2006:426) as evidenced by languages which were perceived to be complex, such as Latin and French, having assumed global language status during their heydays. Crystal (2006:426) elaborates that “a language becomes a world language for extrinsic reasons only, and these all relate to the power of the people who speak it”. This power can be applied in political (military), technological and cultural contexts. Thus the functionality of English spread to several domains to the extent that “so many countries have found it useful to adopt English as a medium of communication, either

for internal or external purposes” (Crystal, 2006:427). The domains are: politics, economics, the press, advertising, broadcasting, motion pictures, popular music, international travel and safety, education and communications.

2.2.2 Attitudes towards English as a world language

The use of English as well as the development and acceptance of localised varieties of the language in the world depends on the attitudes of the host countries. Schmied (1991:144-170) discusses reactions towards English as a world language in terms of different conceptualisations of ‘attitude’. Phillipson (1992), Adegbija (1994), Kachru (1996) and Schneider (2003) highlight the ambivalent nature of the attitudes of the host countries towards English as a global language.

With specific reference to the influence of English on African languages, Schmied (1991:144) observes that the reaction to the use of English elements in African languages differs greatly among Africans, depending on the social characteristics of the community or the individual and the linguistic characteristics of the influence. Schmied (1991:144) notes that the influence of English on African languages or its general use divides the host language-using community into ‘purists’ (those who are resistant to foreign elements and object to their use because they feel their language is being adulterated) and ‘adaptationists’ (those who are accommodating, usually because the influencing language enjoys considerable prestige in the speech community).

In terms of attitude, understood as a general language stereotype, Schmied (1985:247) reports that in many parts of Africa, and indeed in many parts of the third world, English is believed to be ‘beautiful’, ‘rich’, ‘logical’, ‘sophisticated’ ‘pure’, ‘precise’, ‘rhythmical’, ‘refined’, ‘superior’, ‘intimate’ and ‘pleasing to the ear’ indicating that “English seems to enjoy high international prestige as an idealised world language”.

In terms of attitude conceptualised as language beliefs, there are several types of beliefs, usually supported by communicative, national, personal, educational and cognitive reasons. Based on these reasons, there are arguments for and against the use of English in Africa, bringing out a sense of ambivalence in Anglophone Africa. There is also a general feeling that “...one remedy towards improving attitudes towards English, especially among those who associate English with

the colonial past or regard English as an evil influence which necessarily leads to Westernisation, is the recognition of African varieties of English as pedagogical and sociolinguistic models” (Schmied, 1991:171).

Regarding attitudes towards African varieties of English, Schmied (1991:173) notes that, by and large, “arguments against African languages are also used against African varieties of English”. Similar prejudices obtain, and despite the early, and possibly premature christenings of some African Englishes by liberal linguists, the general public appears to prefer Standard varieties of English, particularly British English as a model for African learners of English. However, African varieties of English are tolerated if not actually encouraged at the level of pronunciation.

Phillipson (1992:27) notes the Janus-faced nature of the attitude towards the English language in periphery countries. On one hand, he notes that “the pull towards the English language in periphery countries has been remarkably strong and attitudes towards the language have been favourable” (Phillipson, 1992:27), leading to the displacement and replacement of African indigenous languages. He goes on to give examples of institutions such as the former OAU, now AU abandoning plans to promote the use of African languages at an international arena, citing financial constraints.

However, there are many African countries that were not necessarily Anglophone colonies, e.g. Namibia, Mozambique, Rwanda and Angola, as well as Asian countries like China and Japan which are incrementally ‘appropriating’ the English language. These countries are making frantic efforts to promote the teaching of the English language (Kadenge, 2010; Schneider, 2014). Kadenge (2010) argues that these countries are being compelled to learn English by the need to participate effectively in international affairs, which participation is being conducted in English, by and large.

On the other hand, there are unfavourable attitudes and even opposition to the use of English in other countries. According to Phillipson (1992:35), “those protesting include the colonised people, European parliamentarians, political enemies of core-English nations, guardians of the purity of languages that English intrudes on, and intellectuals from core and periphery-English countries” who all see ‘linguistic imperialism’ in the global advance of English.

Some World English scholars including Pennycook (1998), Phillipson (1992; 2000; 2004; 2008) Skutnabb-Kangas (2000; 2002; 2008), Skutnabb-Kangas and Phillipson (2010) are generally of the view that the global advance of English has actually led to linguistic genocide, the disappearance of indigenous languages in the countries where English has assumed dominance. According to Skutnabb-Kangas and Phillipson (2010:82), “English can be seen as the capitalist neo-imperial language that serves the interests of the corporate world and the governments it influences,” accumulating linguistic capital in the process and dispossessing the indigenous languages of their own linguistic capital.

Thus the title of a paper by Kachru (1996) i.e. “World Englishes: Agony and ecstasy” aptly sums up how geopolitical territories across the world have reacted to the global spread of English. Kachru points out that for some, the implications are agonising while for others they are a cause for ecstasy. In a similar vein, Schneider (2003:233) concurs when he summarises reactions to English as a global language thus:

On the one hand, English is the world’s leading language, the main vehicle of international communication, and in that role it is an essential, indeed indispensable tool for international economy, diplomacy, sciences, the media, and also individual interactions across language boundaries. On the other hand, it has been damned as a ‘killer language’ responsible for the extinction of innumerable indigenous languages, dialects and cultures around the globe.

It may be remarked here that the clear ambivalence in the way English as a global language is viewed evident in this quotation, and indeed in the foregoing discussion, is not just a national stance for a given country but cascades down to the psyche of individual citizens. In other words, a single individual may hate and like English at the same time, love it because it is handy, (“knowing English is like possessing the fabled Aladdin’s lamp...” (Kachru, 1986:1)) and hate it because it is a language brought by foreigners. In section 2.8 below, it is pointed out that the government of Zimbabwe detests everything western except their language (Ndlovu, 2011) because they find the language expedient. A study carried out by Nkwe and Marungudzi (2015) on the use of English as a medium of instruction in Zimbabwean secondary schools also testifies to this ambivalence. It is such background against which the ascription of variety status to ZimE will be explored in Chapter 4.

2.3 Study of world Englishes: Theoretical conceptualisations

As we saw in Chapter 1, the study of World Englishes as an academic discipline is relatively recent and can be traced, as shown in the section below, from Braj Kachru's (1982) pluricentric conceptualisation of global varieties of English and his seminal three concentric circles (the inner circle, the outer circle and the expanding circle) model, through Platt *et al.*'s (1984) new Englishes model, Schneider's (2003; 2007) clearly influential Dynamic Model and Mesthrie's (2008) English Language Complex model, among others. As indicated in the previous chapter, other terms have been used to refer to world Englishes, among them 'new Englishes' (Platt *et al.*, 1984) 'postcolonial Englishes' Schneider (2003; 2007; 2011) and 'global Englishes'. Just to clarify the terminology, Platt *et al.* (1984:2-3) define 'new Englishes' as varieties developed through an education system in which English is a medium of instruction though English may not be the main language used by the inhabitants of a polity, while 'world Englishes' generally entails "institutionalised second language varieties of English spoken around the world" (Kachru *et al.*, 2008:2). From these different conceptualisations, it is Schneider's dynamic model that will be given most attention as it sheds light on the manner in which ZimE has developed.

It is worth noting that Platt *et al.* (1984) take new Englishes as objects worthy of study in their own right and that these correspond to Kachru's (1982) outer circle varieties. What is crucial about Schneider's (2003, 2007) model is that the model successfully unifies the Inner and Outer circle aspects in his STL (settler) and IDG (indigenous) strand, without abandoning the valuable contrasts and conceptual focus introduced by Platt *et al.* (1984) and Kachru (1982).

As indicated in Chapter 1, Schneider (2003; 2007) views new Englishes in different geopolitical zones as sharing the same sociolinguistic and language contact conditions, despite the fact that they have traditionally been studied independently of each other. Thus he adopts a diachronic approach in a bid to come up with a coherent theory that unites new Englishes under a fundamentally uniform developmental process. Schneider (2003:234) argues that the structural and sociolinguistic similarities among the new Englishes, despite the different settings, are a result of a uniform process "determined by general sociolinguistic principles and characterised by a significant set of common traits..." and contact processes that may be explained through reference to theories of communication, accommodation and identity formation. It can also be

pointed out that this model provides useful criteria for the resolution of the identity and status questions of new varieties of English, as will be explained in section 2.6 below.

To explain the origin and evolution of new Englishes, Schneider (2003; 2007) proposes a framework known as the Dynamic Model. This model comprises five developmental phases of new Englishes i.e. the foundation phase, the exonormative phase, the nativisation phase, the endonormative phase and the differentiation phase. From these phases, it is clear that the origin and development of new Englishes is “a process of linguistic convergence, followed by linguistic divergence only later...” (Schneider, 2003:236).

In order to relate this model to the study of English in Zimbabwe, it is imperative to outline the key characteristics of these five stages. The foundation stage is the initial stage when English begins to be used regularly in a new territory (Schneider, 2003:244). This is usually a result of the relocation and subsequent settlement by a significant number of English speakers for an extended period. Such settlement may also be a result of emigration. Contact between the dialects of the settlers, who usually come from different dialect backgrounds, leads to koineization, a process involving speakers of each dialect mutually adjusting their pronunciation and lexical usage to facilitate mutual understanding, resulting in the emergence of a relatively homogeneous middle-of-the-road variety which is of a phonetically and grammatically intermediate nature (Schneider, 2003:244). In turn, there will also be contact between the language of the settlers (the superstrate) and that of the indigenes (substrate). This is a phase that is also marked by restricted contact between the settler and the indigenous groups and marginal bilingualism. In terms of linguistic forms, the indigenous languages do not influence the English spoken by the settlers except for place names that are a result of heavy toponymic borrowing.

The second phase (exonormative stabilisation) usually follows stabilisation of settlements under foreign dominance. Schneider (2003:245) points out that it is a phase during which English is used more regularly in a new environment by expatriate native residents. The British form of English is usually accepted as the standard before adjustment to the local environment soon creeps in, with the consequence that the superstrate strand of English begins to move towards being a local language separate from British English. Schneider (2003:245) also notes that the exonormative stabilisation is courtesy of a self-perception and social identity of the expatriate

residents as outposts of Britain. Though there is a conservative and unaltered cultural and linguistic orientation, there is adoption of local vocabulary, borrowing and new coinages (Schneider, 2003:245-246). From the IDG perspective, this is also the period that the command of English turns into an important asset for commercial purposes, translating into a positive attitude towards English, thus creating a new elite among the indigenous people (Schneider, 2003:246). The exonormative phase is also the stage that marks the beginning of structural nativisation, starting at the spoken level (Schneider, 2003:246). The phase is also marked by processes of code switching, code alternation, passive familiarity, second-language acquisition strategies and negotiation; processes that are also associated with the early stages of pidginisation and creolisation, as observed in section 2.7 below.

The exonormative phase is followed by the nativisation stage (Schneider, 2003:247-249), a stage when linguistic and cultural transformation is most pronounced. It is a period when it dawns on settler speakers that they are no longer under the linguistic and cultural guidance of the mother country (Britain). This stage coincides with political independence in many countries and by this stage indigenous speakers have developed local linguistic idiosyncrasies though this is tempered with the ‘complaint tradition’ from language purists. This period is also characterised by heavy lexical borrowing and use of loan words and a local accent from the indigenous population that is traceable to the phonology of the indigenous languages.

The endonormative stage (Schneider, 2003:249-253) that follows nativisation is marked by a gradual acceptance and adoption of an indigenous language norm and a new locally-rooted self-confidence best captured by such pronouncements as: “In language we can and must go alone” as witnessed in the development of New Zealand English (Schneider, 2003:250). Sometimes, progression to this stage is prompted by some catastrophic political event (called Event X) (Schneider, 2003:250). The local forms of English are now accepted even for formal purposes. Differences between the superstrate strand and the substrate strand dissipate and the resultant language variety is perceived as homogenous even though there could be some heterogeneity (Schneider, 2003:251). By this stage almost all the members of the indigenous group have undergone a process of language shift with the consequence that the original indigenous languages are endangered. The endonormative stage is also characterised by vigorous literary creativity in the new variety and it is the time when the new English is no longer referred to as

‘English in X (X standing for a country or some such geopolitical territory) but as ‘X English’ (X standing for the now ontologically existing new English in a particular country or polity).

The final phase (differentiation) is the turning point (in the sense that it allows for diversification) in the evolution of the new English where the emergence of a new variety is now a thing of the past (Schneider, 2003:253). The inhabitants of the new nation now narrow their identity construction from a national scale to an immediate community scale, where they regard themselves as a composite of subgroups with different ethnicities, regional backgrounds, gender, age and social status, though the earlier sense of national unity does not die out. This phase is also marked by the emergence of new varieties (manifest in accents, lexical expressions and structural patterns) to mark new identities (Schneider, 2003:253) and a shift in focus from the collective identity characteristic of the preceding stage (the endonormative stage) to group identification and social categorisation (Schneider, 2003:253).

Schneider (2003:254) offers the caveat that despite the theoretical neatness in this model, there are fuzzy boundaries between the phases of the dynamic model and that not all characteristics of each phase need to be present for the phase to hold. It is also possible that not all nations go through all the phases; some fossilise at particular stages for various reasons. Most importantly it can be observed throughout Schneider’s description of the dynamic model that for all linguistic changes and realignments that occur in each phase, there is a corresponding realignment on the cultural identity plane, though the role of cultural identity has been discounted by some scholars such as Trudgill (2004).

As we seek to determine the variety status of the English used by Black second language users of English in Zimbabwe, explicit reference to this model will be made in 2.8 below as well as in Chapter 4. This reference is important because as will be shown in 2.6 below, Schneider’s (2003; 2007) dynamic model just described clearly speaks to the determination of the nature and identity of new Englishes by incorporating linguistic and sociolinguistic criteria to determine variety status. The model also shows that a new English status and identity is not a yes or no kind of question but a dynamic state of affairs in which all postcolonial situations contain new Englishes in various stages of development. Moreover, it can be noticed from this discussion that there is an indissoluble bond between the linguistic identity and the sociocultural identity of the

speakers. It is this close relationship that bears on the question of variety acceptance and institutionalisation, as will be illustrated in section 4.4.

However, since its publication and subsequent empirical application and theoretical interrogation, e.g. by Mukherjee and Gries (2009), Tan and Low (2010), Van Rooy and Terblanche (2010), Spencer (2011), Werner (2013), among others, the Dynamic Model has been subjected to close scrutiny. Schneider (2014) summarises some applications and criticisms of the model, among them the rejection by Trudgill (2004) of the role of social identity and attitudinal perspectives in the development of new Englishes. Van Rooy and Terblanche (2010), though agreeing to a great extent with Schneider's proposals have also pointed out that there should be a tempering of the first language acquisition perspective with the second language acquisition perspective in order to enhance the understanding of the development of new Englishes. According to Van Rooy and Terblanche (2010:16), in other words there is need for a more intensive consideration of the input since the extent and availability of the native-speaker input surely affects the development of new Englishes.

Another criticism of the Dynamic Model is the bunching of Inner Circle and Outer Circle varieties in the model basing only on the Asia-Pacific region (Mesthrie & Bhatt, 2008), a point that is, on the contrary, understood as a strength by Van Rooy and Terblanche (2010). There were also concerns that the model is not quite applicable to all new Englishes in the expanding circle (Sergeant, 2012), a point Schneider (2014:16) rebuts, arguing that though the model indeed explicitly relates to postcolonial Englishes, the generalisation to other contexts is actually an intentional component of the model courtesy of the level of granularity he desired. However, he goes on to suggest a modified model for the expanding circle i.e. the 'transnational attraction' model.

From Bolton's (2005) description of research traditions and orientations in the study of world Englishes as outlined below, it can be argued that Schneider's approach is a conglomeration of the English Studies, the sociolinguistic perspective as well as that of linguistic futurology, the last perspective being more apparent particularly where the dynamic model has potential to predict what may become of the new varieties of English spoken in different geopolitical territories, though the question of linguistic futurology has also been questioned.

Despite the criticisms just sketched, Schneider's (2003; 2007) Dynamic Model remains relevant and applicable to the study of new Englishes in the Outer Circle (including ZimE), and in a modified version in the Expanding Circle. In comparison with Platt *et al.*'s (1984) new Englishes model, Mollin's (2007) model, Trudgill's (2004) 'new dialect formation' as well as other ways of determining variety status suggested by researchers on ZimE (e.g. Makoni, 1993; Mlambo, 2009), the Dynamic Model remains appropriate courtesy of its potential to factor in both linguistic and social factors as well as to address different kinds of new Englishes contexts. The model is also dynamic enough to accommodate any new English depending on the cultural and linguistic characteristics present in the new variety. Most importantly, against the backdrop of scanty research on ZimE as a general phenomenon, the Dynamic Model is most appropriate in the sense that it provides a bird's eye view on the new variety with potential for future studies to zoom in on more specific features identified in the trailblazing study and subvarieties or registers of ZimE.

2.4 Research traditions and orientations in the study of world Englishes

There are a number of identifiable research traditions and orientations, as pointed out in Chapter 1 from which new Englishes may be and have been studied. Bolton (2005) identifies eleven such traditions in these studies, spanning from the pidgin and creole studies of the 1930s to the linguistic futurology of the late 1990s and the present. Table 2.1 (adopted from Bolton, 2005: 70 – 71) summarises these traditions. These traditions show the strands of different focuses regarding the study of World Englishes and their contributions and thus demonstrate the progression of the discipline, contextualising and justifying the orientation of the current study.

Table 2.1: Research traditions and orientations on the study of world Englishes

Tradition/ Orientation	Exponents	Objective(s)	Timeline
English studies	Quirk (1962; 1972; 1990), Burchfield (1985), Greenbaum (1985), Quirk and Widdowson (1985), McArthur (1992; 1998), Görlach (1995, 1998), Schneider (1997a; 1997b).	To describe varieties of English from eclectic descriptive and historical perspectives. Situated against a tradition of English studies (Anglistik), dating from the late nineteenth century, e.g. the work of Otto Jespersen, Daniel Jones and Henry Sweet.	1960s – present
English corpus linguistics	Greenbaum (1996), Nelson (1996), Meyer (2002).	To provide accurate and detailed linguistic descriptions of world Englishes	1990 - present
Sociolinguistic approaches (The sociology of language)	Fishman (1972), Fishman, Cooper and Conrad (1977), Fishman <i>et al.</i> (1996).	To conduct research on English in relation to such issues as language maintenance/shift and ethnolinguistic identity.	1960s – present
A features-based approach	Trudgill and Hannah (1982; 1994), Cheshire (1991).	To describe varieties of English through dialectological and variationist methodologies. Situated against the long tradition of British and European dialectology.	1980s – present
Kachruvian studies	Kachru (1982;1983, 1986 etc.), Smith (1981; 1987), Lowenberg (1984), Gupta (1994), Pakir <i>et al.</i> (1994), Bautista (1997), Kirkpatrick (2002), Bolton (2003), Modiano (2003).	To promote a pluricentric approach to world Englishes, highlighting both the ‘sociolinguistic realities’ and ‘bilingual creativity’ of Outer-Circle (and Expanding-Circle) societies.	1980s – present
Pidgin and Creole studies	Reinecke (1937), Todd (1984), Mühlhäusler (1986), Romaine (1988), Holm (1988/1989), Sebba (1997), Mufwene (2001).	To describe and analyse ‘mixed’ languages and the dynamics of linguistic hybridization (beginning with the early work of Hugo Schuchardt 1842 – 1927)	1930s – present
Applied linguistics	Halliday, MacIntosh and Strevens (1964), Strevens (1977; 1980), Brumfit (1982), Kachru (1982; 1986; 1992a).	To explore the implications of world Englishes for language learning and teaching.	1960s – present

Lexicography	Webster (1806; 1828) [USA], Yule and Burnell ([1886] 1969) [India], Butler (1981) [Australia], Orsman (1997) [New Zealand], Allsopp (1996) [Jamaica].	To fulfil a range of objectives, including the expression of a national linguistic identity.	1980s – present (within the world Englishes context)
Popularisers	McCrum, Cran and McNeil (1986), Crystal (1995; 1997).	To interest a mass reading public in issues related to world Englishes.	1980s – 1990s
Critical linguistics	Phillipson (1992; 2000, 2003), Pennycook (1994; 1998).	To express resistance to the linguistic imperialism and cultural hegemony of English. Derived from a Marxian political analysis and/or postcolonial theory.	1990s – present
Linguistic futurology	Graddol (1997)	To predict future trends in the spread of English and English language teaching worldwide.	1997 – present

In order to contextualise the study of the morphosyntactic features of ZimE using the corpus linguistics method in terms of research orientation and indeed other research studies reported in sections 2.7, 2.8 and 2.9, Table 2.1 has given a comprehensible outline of all the traditions and orientations in the study of world Englishes identified by Bolton (2005). However as noted earlier, the present study falls into the English corpus linguistics and to some extent the Kachruvian studies traditions as it is more concerned with the linguistic characteristics and sociolinguistic status of a specific variety of English. The corpus-based method, which was used in the collection and analysis of data will be described in more detail in Chapter 3.

As we saw in the previous chapter, the term ‘world Englishes’, which has come to be used interchangeably with other terms such as new Englishes was first used by Kachru (1982; 1985; 1986; 1988; 1990; 1992 etc). Kachru (arguably the father of World Englishes as an academic discipline) used ‘world Englishes’ to refer to the pluricentric and distinctive varieties of the English language spoken in different parts of the world. However recent sentiments, e.g., Mesthrie and Bhatt (2008:3) indicate that it is problematic to use these terms. Mesthrie (2006) as well as Mesthrie and Bhatt (2008) go on to suggest using the term ‘the English language complex’ (ELC) instead. Similarly, other scholars, e.g. James (2008), feel that it is no longer

tenable to view the current varieties of English as territory-specific, but to view them as lingua francas used in unique ways by different players on the international market.

Kachru's initial approach to the study of world Englishes was to classify the Englishes into clusters which he called circles. (See Kachru, 1982 and related articles). It is important to note that Kachru's (1982) conceptualisation of world Englishes came largely as a response to earlier approaches (dominant among them the deficit approach and the deviational approach) that viewed varieties of English in the Outer Circle as deserving no recognition as English varieties (Kachru, 1990). Kachru disagrees with Quirk (1987; 1988), one proponent of the deficit approach. In Quirk's view, the recognition of a sociolinguistic identity for varieties of English in the Outer Circle is a result of the "false extrapolation of English 'varieties' by some linguists." Quirk goes on to reject the use of identificational terms such as 'Nigerian English', 'West African English', 'South Asian English', 'Singapore English' and so on characterising them as misleading if not entirely false. Quirk (1991:20) notes that the problem with 'liberation linguistics' as he dubs Kachru's views is that it undervalues "the baby of Standard English while overvaluing the undoubtedly important bathwater of regional, social and ethnic varieties," something which is "neither liberal nor liberating" as it permits "learners to settle for lower standards than the best and it is a travesty of liberalism to tolerate low standards which will lock the least fortunate into the least rewarding careers [sic]". It is such views and a host of other sacred cows of English (i.e. acquisitional sacred cows, sociolinguistic sacred cows, pedagogical sacred cows, and theoretical sacred cows) that Kachru (1988) proposes should die if we have to come up with insightful frameworks in the study of world Englishes.

From a different perspective, Bruthiaux (2003) does not take issue with the term 'World Englishes', but criticises the three-circles model for not having much explanatory power because "it is descriptively and analytically inconsistent as well as over-representative of a political agenda". According to Bruthiaux (2003:161), the three concentric circles model has helped to assign more value to otherwise denigrated varieties by drawing attention to similarities with the older, more established varieties, and by challenging perceptions that they are not communicatively as useful and lack prestige. However, he adds that the Kachru model still has a number of limitations: it suffers from being based on a political/historical view of English worldwide, and hence fails to capture transplantations of the language in areas that have not been

colonised formally. It also ignores local variation by encouraging broad-brush descriptions of the manifestations of English worldwide, without performing a more specific and detailed sociolinguistic analysis.

According to Xiao (2009) world Englishes have been studied from two broad perspectives; the sociocultural perspective (to which Kachru and Schneider's approaches belong) which is concerned with the elaboration of theories and models of development, spread, classification and interaction of new Englishes; and the linguistic perspective which focuses on the linguistic features in selected varieties of English at various levels including phonetics and phonology, morphology, syntax and discourse. This view concurs with Mesthrie's (2003) approach i.e. "establishing a truly comparative data-base for linguistic analysis" as noted earlier. For the purposes of this study, it is by and large, the linguistic approach that has been adopted. The variety of English that is used in Zimbabwe as represented in both oral and written texts produced by Black second language users will be analysed for morphosyntactic features and its variety status subsequently determined.

2.5 The linguistic tradition in the study of world Englishes

Though meaningful contributions to the study of world Englishes can be noted from the traditions outlined above, it can be observed, as indicated in the previous chapter, that a more effective way of ascertaining the status of a new English variety is through scientific methods such as the corpus approach, which is based on the linguistic features of the variety as actually used. Thus the linguistic tradition may be viewed as complementing the sociolinguistic tradition that focusses on the broad functions of a language or language variety within a polity. In carrying out a linguistic analysis of a variety, a number of linguistic features are usually considered. These features, according to Kortmann *et al.* (2004), Lunkenheimer and Kortmann (2011) and Kortmann and Lunkenheimer (2012) belong to the following broad categories and yield about 235 narrower features:

- Pronoun
- Noun phrase

- Verb phrase: tense and aspect
- Verb phrase: Modal verbs
- Verb phrase: Verb morphology
- Adverbs
- Negation
- Agreement
- Relativization
- Complementation
- Discourse organization and word order

Studies that have already been carried out (see review of such studies in sections 2.8, 2.9 and 2.10 below) have yielded the result that new varieties of English depart from the so-called standard forms in the aspects listed above in systematic ways. The present study hopes to analyse the corpus of ZimE and determine the extent to which the morphosyntactic features present therein confirms these findings. As will be emphasised in 2.6 below, the determination of the linguistic characteristics of new varieties are a critical cog in resolving the variety status question.

One notable global study that has used the linguistic feature approach is the *electronic World Atlas of Varieties of English* (eWAVE) which was designed and compiled between 2008 and 2011. eWAVE reports on morphosyntactic variation in spontaneous spoken English in 74 varieties of English in different parts of the world that include Africa, Asia, Australia, British Isles, Caribbean, North America, Pacific, and the South Atlantic. One strength of eWAVE is that it is an open access resource that is reasonably comprehensive in terms of the number of varieties covered. In 2012 De Gruyter Mouton also published the *Mouton World Atlas of Variation of English*, which offers accounts of the individual 74 data sets in eWAVE as well as large-scale comparisons and synopses across the individual variety types and Anglophone world regions. Lunkenheimer and Kortmann (2011; 2012) argue that eWAVE thus has the potential for serving both as a teaching tool in academic teaching around the world and as an indispensable research tool for specialists in many different fields of linguistics including the study of world Englishes.

In terms of research, it can also be observed that, as Schneider (2007:16) points out, the earlier traditions tended to focus on specific regions or countries, and therefore the comparative perspective and broader patterns did not receive enough attention. eWAVE and corpus linguistics (which is used in the current study) are useful tools to achieve better comparability between varieties. According to Lunkenheimer and Kortmann (2011), eWAVE also facilitates the investigation of global-scale patterns of morphosyntactic variation in English and helps in answering questions like the following:

- Which features are most/least widespread across varieties of English worldwide?
- How many varieties of English worldwide share feature X?
- Is feature X restricted to or characteristic of a particular part of the English-speaking world?
- Is feature X restricted to or characteristic of a particular group of varieties?
- Does variety A have feature X?
- In which area of grammar does variety A differ most from variety B?

Lunkenheimer and Kortmann (2012) go on to suggest that the information provided in eWAVE can also be used for the investigation of more general questions, such as the following:

- Which features generally are characteristic of a particular variety type (e.g. L2 varieties)?
- In which domain of grammar is there most/least heterogeneity/homogeneity among varieties of English worldwide?
- Are English-based pidgins and creoles, as a group, significantly different from other varieties in terms of morphosyntax?

A project similar to eWAVE is the MPI-EVA's APiCS (Atlas of Pidgin and Creole Language Structures) which appeared in 2012, both as a book atlas and as an electronic database like eWAVE.

It can be noted from these descriptions that the eWAVE, APiCS and the *Mouton World atlas of variation in English* (MWAVE) are both an extension of the sociolinguistic perspectives of Schneider (2003; 2007) and Mesthrie (2003) to accommodate the linguistic features of the varieties of English being explored. As pointed out earlier, the present study takes a similar

approach and engages some of the questions (outlined above) that these projects can potentially address. However, though these two projects have extended the sociolinguistic traditions and delved into the linguistic aspects of the varieties, they still relied to some extent on anecdotal data and conjecture, making it important for current research to turn to corpus methods, which are potentially more reliable.

2.6 Determining statuses of new Englishes

Mechanisms for ascribing variety status to new Englishes are not always straightforward. I believe that an effective way of determining the variety status of a new English is to complement sociolinguistic approaches with the linguistic approach. The linguistic approach however has to be able to separate actual innovations from errors. A number of studies have attempted to explain the differences between the two. In the words of Van Rooy (2011:189), “a distinction between error and conventionalised innovation is essential to understanding if and how new varieties of English develop new conventions”. This distinction is useful because it is what bears on whether or not to confer variety status on particular varieties of English.

Though not of a purely theoretical nature, the study by Van Rooy (2011) suggests that two important criteria may be considered in determining whether a linguistic form is an error or an innovation i.e. grammatical stability and acceptability. Using corpus methods to study the use of the progressive aspect (BE + -ING), ‘can be able to’ and ‘enable + bare infinitive’ Van Rooy (2011) comes to the conclusion that new linguistic conventions evolve from forms that may have started out as errors. The linguistic features to be discussed in Chapter 4 will be analysed in terms of whether they are errors or innovations and in terms of where they are coming from an attempt to highlight the process of the development of linguistic innovations in world Englishes.

Closely related to Van Rooy (2011) is a study by Gut (2011). Interrogating the question as to whether structural innovations found in new Englishes may be regarded as errors or transfer phenomena, Gut (2011) makes the observation that the question may only be resolved through the new English users’ attitudes towards the new English and through whether the variety has an internal norm orientation or an external one. Varieties which command positive attitudes and

evince internal norm-orientation usually pass the conventional or institutionalised new English test. Thus, as I mention elsewhere the variety status question can only be conclusively settled through an appeal to both linguistic and sociolinguistic considerations.

Focussing on whether linguistic innovations are an index of a new variety or simply learner English and using the case of Euro-English, Mollin (2007) postulates that an important criterion for deciding the status of ESL varieties should be based on the three processes of expansion in function, nativisation of form and institutionalisation of a new standard. Two of the processes i.e. expansion and institutionalisation would fall in the sociolinguistic orientation while nativisation would be linguistic in orientation. A consideration of varieties using this criterion brings out the fact that ascribing variety status is really a complex undertaking, a point well-made by de Klerk (1999) in the context of South Africa. However the inclusion of the linguistic dimension in Mollin's (2007) criteria resembles a part of Schneider's (2003; 2007) Dynamic Model and is quite promising. In resolving the variety status of ZimE, Mollin's (2007) criteria will also be harnessed.

On the other hand, Bamgbose (1998:2) views an innovation as "an acceptable variant, while an error is simply a mistake or uneducated usage". Bamgbose (1998:2) adds that if innovations are seen as errors, non-native varieties of English may never be recognised as legitimate varieties. Bamgbose (1998:2) makes the point that "there is often an indeterminacy between what counts as an innovation and what is to be regarded as an error" and suggests two ways of resolving the problem: appeal to an external standard and appeal to internal factors. Bamgbose reaches the conclusion that appealing to an external standard is not quite useful since there are a number of levels of the external norm types; the code norm, the feature norm and the behavioural norm and apart from that there is the argument why a native-based standard should continue to license the norms of non-native Englishes (Bamgbose, 1998: 3).

As for the appeal to internal factors Bamgbose (1998) lists five factors that must be considered in deciding the status of an innovation: demographic, geographical, authoritative, codification and acceptability factors. Demographically the bigger the number of acrolectal users of an innovation, the higher the chances that the form may be accepted as an innovation. Geographically, the greater the geographical spread of an innovation the higher the chances of its

acceptance as an innovation. In terms of codification, once an innovation enters a dictionary as correct, its legitimacy as a regular form is assured.

Closely related to the issue of codification is one of authority which has to do with the actual use of an innovation in various sectors of life (Bamgbose, 1998:4). Though the source of authority may not be difficult to identify in native varieties, establishing the source of authority in non-native varieties is inevitably difficult especially in the absence of codification. Finally, in terms of acceptability, once an innovation is accepted, it would have passed the final hurdle so it is assured of a reasonable lifespan. Bamgbose (1998:4-5) observes that acceptance and codification are the most critical factors in the determination of the status of a linguistic feature as an error or innovation.

According to Platt and Weber (1980) criteria for ascribing new English status include: the development of the variety through the educational system; use of the variety in a range of purposes such as letter writing, the writing of creative literature, parliamentary debates and media reporting; and the exhibition of 'localised' features in pronunciation, sentence structure and the lexicon. As Makoni (1993:100) rightly points out, the criteria for ascribing variety status are therefore both linguistic and sociolinguistic.

A discussion of the Dynamic Model (Schneider, 2003; 2007) in section 2.3 above has brought to the fore the universal applicability of the model to different new Englishes. It was noted that one of the strengths of the Dynamic Model is its incorporation of the linguistic and cultural/attitudinal factors in the determination of new English variety status. In this regard there is a big area of overlap between Schneider's model and both Mollin's (2007) and Platt and Weber's (1980) models. According to Schneider (2003) new Englishes are born (foundation), develop, mature (exonormative stabilisation and nativisation), solidify (endonormative stabilisation) and split (differentiation) alongside cultural identity and are therefore in a constant state of flux where it is possible to put a developmental tag on English used in any postcolonial setup.

The foregoing discussion shows, among other issues, that indeed the question of the status of varieties of English strongly hinges on whether the linguistic features it evinces are innovations or errors. This strengthens the case for the adoption of the linguistic approach taken in this study,

more so in the context of ZimE as discussed in 2.10 below. It emerges from 2.10 below that though the researchers are exactly right on the sociolinguistic functions of English, their observations on the structural features of ZimE are largely based on introspective conjecture. The linguistic approach adopted in the present study in which the morphosyntactic features will be described will then be complemented by a more analytical discussion of the forces behind the features. Thus, it emerges from this discussion that the variety status question cannot be resolved without exploring the linguistic features of a new variety. Actually, it is only when the linguistic features have been established that social issues as raised by Platt and Weber (1980), Bamgbose (1998), Mollin (2007) will be settled. It is against these linguistic and sociocultural criteria that the variety status question will be addressed in Chapter 4 subsection 4.

2.7 New Englishes and pidgins and creoles

In as much as “native Englishes, indigenised Englishes and English pidgins and creoles have all developed by the same natural restructuring processes” (Mufwene, 2001:223), it is important to interrogate the relationships between pidgins and creoles on one hand and new Englishes on the other. In fact, the new Englishes discipline has its parentage in the study of pidgins and creoles (Mesthrie & Bhatt, 2008). The rationale for this discussion is that, as indicated in 2.1 above, certain linguistic features of new Englishes may actually be explained through reference to processes of pidginisation and creolisation since these processes have a lot in common with new Englishes given that they both emerge out of contact situations. The discussion also describes the two processes using detail against which the occurrence of particular features of ZimE may be explained.

Despite the commonalities, it is however, important to differentiate these two phenomena. A pidgin refers to a language that is nobody's native language (Wardhaugh, 2006:61) and which usually arises when two speakers of different languages with no common language try to have a makeshift conversation. Pidgins are languages that are born after contact between at least two languages (Kirkpatrick, 2007:13; Jenkins, 2009:11). Kirkpatrick (2007:13) goes on to note that many pidgins developed during the period of empire and international trade with one of the ‘parents’ being a European language such as English or French and the other parent being the

language of the people the Europeans were trading with. Kirkpatrick (2007:13) also observes that between the two languages, “one provided the vocabulary items and the other one provided the grammatical structure”.

The word ‘pidgin’ comes from a Greek word referring to business, indicating that pidgins largely originated in business contexts. The lexicon of a pidgin usually comes from one language, structure often from the other. Because of colonialism, slavery etc. the prestige of pidgin languages is very low. Many pidgins are ‘contact vernaculars’ and may only exist for one speech event (Jenkins, 2009:10). Typical examples are Chiraparapa in Zimbabwe, Fanagalo in South Africa and Cikabanga in Zambia.

Jenkins (2009:11) notes that some of the important characteristics of pidgins are that they have no native speakers, they are social rather than individual solutions, are characterised by norms of acceptability and usually develop from simpler to more complex systems as communicative requirements become more demanding. Because they respond to unique and ever-changing communicative requirements of their speakers, pidgins are constantly in a state of flux; thus a pidgin may begin as a jargon pidgin, develop into a stable pidgin and eventually into an expanded pidgin.

On the other hand, when a pidgin is learnt as a mother tongue, it is known as a creole (Kirkpatrick, 2007:13) therefore a creole may be defined as a language that was originally a pidgin but has become a native language to a community of speakers. The term has also been used to designate the language(s) of people of Caribbean and African descent in colonial and ex-colonial countries such as Jamaica, Haiti, Mauritius, Réunion, Hawaii, Pitcairn, etc (Jenkins 2009:62). It is important to note that a creole is a speech community’s first language (Jenkins, 2009:11).

There are a number of theories pertaining to the origin of pidgins; including the independent parallel development theory, the nautical jargon theory, the theory of monogenesis and relexification, and the baby-talk theory (Jenkins, 2009:13-14). However, these theories will not be detailed here.

In terms of grammatical structure, English-based pidgins are characterised by the following:

- few inflections in nouns, pronouns, verbs and adjectives such as absence of number and gender in nouns, absence of case marking in pronouns resulting in use of ‘I’ and ‘me’ interchangeably;
- use of ‘no’ for negation as in *‘I no tu had’*; and
- absence of complicated constructions such as relative clauses. (Jenkins, 2009:63)

As far as the relationship between pidgins and creoles on one hand and new Englishes on the other is concerned, it should be noted that both phenomena are a result of similar processes of sociolinguistic contact between people from different speech communities; pidgins and creoles mainly through slavery, European settlement, labour migration and war and new Englishes through language spread (a process of language imposition by a greater colonial power whereby the uses or the users of a language increase) especially during the era of imperialism. Important characteristics of new Englishes include the following:

- They have developed in an area where English is not spoken by a large number of people.
- They have developed through the education system rather than being acquired initially as a language at home.
- They are used for a range of functions e.g. as intergroup language, in parliament, in official communication, in the media etc.
- They have become indigenised by adopting words from local culture and nativised by stabilising some structural features associated with local languages and/or the language learning process. (Erling, 2006: 405)

It is these nativised features of the new Englishes if any, that, like other enterprises and publications cited above, the present study seeks to confirm, describe and explicate in the context of ZimE.

Platt *et al.* (1984) suggest that new Englishes are made up of a continuum of sub-varieties, with, on one end new Englishes close to L1 English in terms of structure and pragmatics; and on the other, varieties that show the greatest divergence from L1 norms and they have suggested the use of ‘acrolect’ for those closest to the native variety, ‘basilect’ for those furthest and ‘mesolect’ for those in the middle. This classification presents challenges to attempts at assigning variety status

to new Englishes since it almost precludes the conceptualisation of any new English as a homogeneous entity though it presents opportunities to make comparisons in terms of the nature and extent of nativisation between the sub-varieties.

In terms of attitudes towards new Englishes, Kachru (1992:60) observes a case of linguistic schizophrenia, arguing that within the speakers of non-native communities i.e. “the non-native speakers themselves have not been able to accept what may be termed the ecological validity of their nativised or local Englishes”. One would have expected such acceptance, given the acculturation and linguistic nativisation of the new varieties. On the other hand, the native models (such as RP or General American) are accepted without reservations.” As explained in section 2.3 above, the attitudes of the host communities of new Englishes have a bearing on the ascription of variety to new Englishes. This point will be taken up in Chapter 4 where attempts are made to determine the variety status of ZimE.

It emerges from this discussion that there are commonalities between pidgins and creoles on one hand and new Englishes on the other. Though “no Pidgin English exists in Southern or East Africa” (Mesthrie, 2004:962) it is anticipated that the explication of the morphosyntactic features of ZimE which will be dealt with in Chapter 4 will benefit from a consideration of both the commonalities and disparities between these two phenomena.

2.8 African varieties of English: West Africa, East Africa and Southern Africa

ZimE is a type of African English and it is therefore prudent to review the theoretical, historical and research literature on African varieties of English in general. This discussion thus provides a background to the more specific issue of the morphosyntactic features and variety status of ZimE describing the broader context of the variety.

Focussing on the history of English in Africa, Angogo and Hancock (1980: 66) chronicle the arrival of English in Africa in 1530 and its subsequent development till it became the official language of a considerable number of countries on the African continent. Because it is widely

used in Africa, yet is native to nobody there, determining the form of English that is acceptable was always controversial (Angogo & Hancock, 1980:66).

In a similar vein, Mesthrie (2006:382) argues that because the English language has always “existed amidst a multilingual ethos in which language contact has been ever-present”, it is necessary to merge the history of Englishes because “a monolingual history of English is an oversimplification.” Mesthrie (2006) goes on to argue that England is “only one of the many stopping places for a language that started out in Germanic-speaking parts of continental Europe and now has spread almost all over the planet”. Abandoning the use of the term ‘English Language Family’, as already mentioned Mesthrie views the types of English that constitute English as the ‘English Language Complex’ comprising L1 English (also known as English as a Native Language or as Inner Circle Varieties), Pidgin Englishes, Creole Englishes, English as a Second Language, Immigrant Englishes, English as a Foreign Language, Language Shift Englishes, Jargon Englishes and Hybrid Englishes. Thus indeed, African Varieties of English (AVEs) are “clearly emerging as new personalities in the kaleidoscope of World English” (Angogo & Hancock, 1980:93).

Addressing the question as to whether English is emerging into a single standard or diverging into localised regionalisms, Angogo and Hancock (1980:93) conclude that, in terms of its form, there are identifiable linguistic features that co-occur in all AVEs – a term used by the authors to refer to the variety of English spoken in the specific countries of Africa by Africans and not as natively spoken by the British. Following Banjo (1971), Brosnahan (1958) and Criper (1971) in their classification of the types of English used in Africa, Angogo and Hancock (1980:71), categorise the varieties of Englishes spoken in West and East Africa into four types that correspond to the speakers’ level of education and context of acquisition. The types are:

1. Native English of African-born whites and expatriates
2. Native English of locally born Africans
3. Non-native English spoken fluently as a second language (in several styles)
4. Non-native English spoken imperfectly, as a foreign language (in several styles)

Using Platt *et al.*’s (1984) more familiar terminology i.e. ‘acrolect’ ‘mesolect’ and ‘basilect’, there does not seem to be a neat correspondence with Angogo and Hancock’s taxonomy though

it would seem that Type 3 (Non-native English spoken fluently as a second language), which the authors take as the most prominent if not prototypical AVE corresponds to the acrolect and Type 4 to the basilect. The authors go on to note that like other regional varieties of English, AVEs are distinguished by their phonology, grammar, and lexico-semantic characteristics, with phonology perhaps being the most distinctive.

In terms of the phonology of AVEs, Angogo and Hancock (1980) observe that AVE phonological features are largely influenced by the local African substrates though other features may result from non-African native varieties. The authors argue that it is at this phonological level that AVEs share common features. For example, “even though a Yoruba speaker of Nigerian Vernacular English might say ‘sheap shicken’ instead of ‘cheap chicken’ while an Idoma speaker might not, or a Gambian will say ‘bot’ for ‘but’ when a Ghanaian will say ‘bet’, it does not mean that there are many West African Vernacular Englishes” (Angogo & Hancock, 1980:74-75). Angogo and Hancock (1980:75) observe that there is a phonological cohesiveness to all these kinds of English which identifies them as African rather than, say, West Indian English, and specifically as West African rather than East African

In terms of grammar, AVEs have been seen to reflect African linguistic behaviours in many ways e.g. non-occurrence of ‘more’ in both WAVE and EAVE as well as the lack of distinction between ‘they hurt each other’ and ‘they hurt themselves’. Hickey (2012) and Kortmann and Lunkenheimer (2012) make a similar point and go on to identify these features, among them the use of double determiners (Feature 59), different mass/count noun distinctions resulting in use of plural forms in cases where Standard English has singular (Feature 55), *come*-based future/ingressive markers (Feature 116) and *too/too much/very much* for *very* as qualifier (Feature 222). There are similarly features that do not occur in African varieties of English at all and these have also been proposed to negatively define the varieties. These features, which are attested in all other world regions except Africa, include *she/her* used for inanimate things (F.1), associative plural marked by other elements (F.53) and *was/were* generalisation (F.180).

Lexically, AVEs are characterised by vocabulary adapted to its environment with words such as ‘chop’, ‘dash’, ‘palaver’, ‘coaster’ or ‘coastian’, ‘fridgeful’, ‘upstairs’ (intoxicated), ‘cheat’ (lie), ‘safari’ etc. as examples. AVEs have also been seen to be characterised by a penchant for ‘high

style' and innovative analogies. Semantically, there are also carryovers from the mother tongues of the AVE speakers and these include words such as 'sorry', 'how now', 'try', 'are you alright?' whose meanings depart from their usual meanings when used by native speakers of English (Angogo & Hancock, 1980:76-77).

In addressing the issue of legitimacy of AVEs, Angogo and Hancock (1980) criticise the view that 'the only legitimate dialects of English are those having native speakers, arguing that such attitudes are based on the fallacy that "non-native speakers have no right to manipulate the language". Thus they, in the Kachruvian manner, go on to advocate recognition of AVEs as legitimate varieties of World English.

There are other researchers that have set out to explore the more specific features of particular varieties in East and West Africa. Focussing on the forms of English spoken in parts of Africa and referring to it as BAfE (for Black African English), Schmied (1996) identifies three forms of English based on pronunciation. These are Bantu English (characterised by simplification of the complex vowel system of English), Regional English (characterised by region-specific pronunciation peculiarities) and ethnically-related English (characterised by those pronunciation features that are specific to a particular ethnic group).

In terms of vocabulary, Schmied (1996) observes that the lexicon is particularly prone to regional differentiation with the political field significantly contributing to nation-specific vocabulary. Parliament, political parties or informers or the 'special branch' usually attracted local names such as 'life president' (Kamuzu), 'multiparty' (democracy), 'freedom' (freedom from market taxes) etc. in Malawi were seen as the favourite candidates for Africanised expressions. Foods, local garments and cultural rites are also observed to be other fields characterised by African expressions in terms of vocabulary.

In terms of grammar, there are also instances of simplification and these have been witnessed in different regions of Africa e.g., noun classes are not always marked for number and case, prepositions are used differently, inflectional endings of verbs are 'confused' in complex cases, complex tenses tend to be avoided, the use of continuous forms tend to be expanded to stative verbs, verb complementation varies freely and there are problems with idiomaticity. Subsequent studies elsewhere on the continent as shown in Chapter 4, have since confirmed most of these

findings, working with specific features in specific contexts. For example, in a study intended to look for evidence for the existence of specific grammatical features characteristic of Kenyan English and using 26 grammatical features, Buregeya (2006) found out that about 14 (54%) of these features were deemed acceptable by 60% of the respondents, confirming to a considerable extent the existence of a definitely Kenyan variety of English, at least for the written form.

In the South African context, a number of researchers, among them, Coetzee-Van Rooy and Van Rooy (2005), De Kadt (1993; 2002), De Klerk (1999; 2003; 2004; 2005), Kruger and Van Rooy (2016), Makalela (2004), Meierkord (2012), Mesthrie (1997; 2002; 2004; 2008; 2012), Minow (2010), Siebers (2007; 2013), Van Rooy (2006; 2008; 2011) and Van Rooy and Kruger (2016) acknowledge the existence of Black South African English and some of them go further to analyse the use of specific linguistic features of English in South Africa. It is clear that there exists far more studies on South African English in its various sub-varieties in South Africa than is the case in Zimbabwe, with most of the research focussing on Black South African English. However, I will sample only those studies that focus on the linguistic features and variety status of Black South African English, which is a peer to ZimE as used by Black second language users.

One of the earliest studies that focussed on the variety status of Black South African English was carried out by de Kadt (1993). In her study, de Kadt (1993:314) poses the question whether Black South African English is a non-standard variety of South African English characterised by fossilised features or a learner language continuum or a number of ethnic varieties. According to Bernsten (2001:228) another difficulty in describing Black South African English concern the level of proficiency (from the basilect, the mesolect and the acrolect, discussed earlier) to be used as representative of the variety. This question comes against the background that new Englishes do not make one monolithic variety (Platt *et al.*, 1984; Schmied, 1991). It is clear from this discussion that determining the status of a new English touches on a number of questions that need to be answered before the status can be determined. This has important implications for the criteria used in the present study to determine the variety status of ZimE.

Another study that focusses on the variety status of Black South African English is de Klerk's (1999) article entitled 'Black South African English: Where to from here?'. Aimed at exploring

problems associated with defining this new variety of English, the paper argues that problems with defining Black South African English include whether it is a new variety or just a dialect, whether all levels of competence should be considered as part of the variety, whether it should be considered a spoken or written variety, number of speakers that use it, whether it is owned by only those who have had formal schooling or by anyone who has had a ‘smattering’ of English in informal contexts as well as the presence of code-switching. Thus de Klerk is of the opinion that issues of ownership as well as the meaning of ‘new’ in ‘new English’ are key issues in resolving the variety status of new Englishes. De Klerk (1999:311) concedes though that Black South African English fits Platt *et al.*’s (1984) description of ‘new Englishes’ “very nicely”. Though de Klerk (1999) does not use a purely linguistic tradition in her paper, she raises important sociolinguistic criteria applicable to the assessment of the status of ZimE.

Among other issues such as the comprehensibility and labels of South African English varieties, Coetzee-Van Rooy and Van Rooy (2005) also focus on the variety status of South African varieties. They come to a conclusion that South Africans (particularly Black South Africans) label varieties used in South Africa from ethnic, racial and national perspectives, a practice that had come under criticism from linguists. This shows that there is recognition of a specifically South African variety of English which in turn subdivides into racial and ethnic varieties, at least from the perspective of ordinary non-linguist community members. This study illustrates that an alternative mechanism to determine the variety status of a new English is to canvass the perceptions of the ordinary community.

An important contribution that tackles the question of the status of Black South African English from a linguistic perspective is a synthesis of research articles that focus on the linguistic features of Black South African English made by Makalela (2004). In this paper, Makalela (2004:355-356) undertakes to prove the distinctiveness of Black South African English by providing “an in-depth analysis of its selected features”. An important insight that comes from Makalela’s article is the employment of the linguistic criteria to determine variety status. The features that Makalela reports on, i.e. extension of the progressive aspect to stative verbs, tense sequencing, topic promotion devices, and modality markers, have been studied by other researchers and are discussed below.

Studies that focus on the linguistic aspects include that by Mesthrie (2004) which summarises the morphology and syntax of Black South African English and finds out that Black South African English mainly differs from other varieties of South African English in terms of discourse organisation and phonetics and not so much in grammar. Though not very pronounced, grammatical patterns characteristic of Black South African English that Mesthrie (2004) identifies include variations in combinations of tense and aspect, subject-verb agreement problems, problems with negation, relativisation, complementation, word order, and pronominal systems. This study also indicates that Black South African English shares many syntactic features with Sub-Saharan peer varieties, more so with East African varieties than West African ones. Mesthrie (2012) also participated in the compilation of *The Mouton World Atlas of Variation in English* and reported related findings about Black South African English. These are interesting findings which open possible avenues for comparison with the Zimbabwean variety that this study focusses on.

In her 2003 article and using a Xhosa English corpus, De Klerk postulates that the variety of English used by Black South Africans should not be studied only as a second language from an applied linguistic point of view, but should also be studied as a language in its own right. In 2005, De Klerk proceeds to single out the use of *actually* in Xhosa English and analyse how it is used by Xhosa speakers of English in order to provide a basis for comparison between Xhosa English and mother tongue English norms. She finds out that speakers of Xhosa English use *actually* three times more than mother tongue English speakers and that they use it more often as a contemplative vocal filler as well as a connective to initiate a turn. In a paper that focussed on the pragmatic use of *well*, De Klerk (2004) also proved the uniqueness of Xhosa English use of the construction. These studies demonstrate the distinctiveness of Xhosa English, a sub-variety of Black South African English.

Another study that takes a linguistic perspective is Van Rooy (2006). Focussing on the extension of the progressive aspect in Black South African English and comparing it to inner circle English and German learner English, Van Rooy (2006) concludes that most uses of the progressive aspect by Black South African English speakers are consistent with conventions in inner circle varieties though he reports significant differences e.g. Black South African English speakers' data display a kind of continuous aspect without temporal immediacy. Van Rooy (2006) argues

here that this could be a result of influence from the persistive aspect of the Bantu languages. In a related study Van Rooy and Kruger (2016) argue that the extension of the progressive aspect becomes entrenched as a conventionalised feature of Black South African English as a function of a fragile balance between sufficient presence in the written and spoken language of learners and a saliency not sufficient enough to entail normative intervention, a point elaborated in greater detail in Kruger and Van Rooy (2016).

Against this background, a review of AVEs, particularly their linguistic aspects, provides a background that can potentially illuminate the occurrence of particular morphosyntactic features in ZimE, given that ZimE is like a peer to those AVEs. It also emerges from the foregoing discussion that considerable research on AVEs has focussed on the phonological and lexical level of linguistic analysis, making it necessary to complement findings in these areas with linguistic features of a morphosyntactic nature.

Having surveyed the phenomenon of new Englishes in different parts of Africa and discussed some of the characteristics of these Englishes that have been identified by researchers and having observed that ascribing variety status to new Englishes depends on a number of factors, both linguistic and non-linguistic, issues that have implications on the mechanisms for determining the variety status of ZimE, it is now appropriate to turn our attention to the Zimbabwean context, the focus context of the present study, and give a historical background relating to the sociolinguistic configuration of the country as well as preview some studies that have been carried out on ZimE. Sections 2.9 and 2.10 respectively address these issues.

2.9 The sociolinguistic history and status of English in Zimbabwe

As indicated in Chapter 1, English came to Zimbabwe through the colonial expansion of the erstwhile British Empire. According to Schmied (1996), colonial penetration into the former Central Africa (Zimbabwe, Zambia and Malawi) in the middle of the 19th century brought the Africans under the influence of missionaries (especially in Malawi), miners (especially in Zambia) and settlers (especially in Zimbabwe). The territory north of South Africa had of course already been explored by hunters, prospectors and missionaries earlier (Hole, 1936; Raftopoulos

& Mlambo, 2009). It was only in 1888 that Lobengula, the king of the Ndebele gave mining rights to Cecil John Rhodes. In 1890, with the blessing of the British government, Rhodes marched into present-day Zimbabwe with the 'Pioneer Column' consisting of two hundred white settlers, some black servants and 500 soldiers and permanently settled in present-day Zimbabwe. It can therefore be suggested that this year could mark the beginning of the foundation phase in terms of Schneider's dynamic model (see Schneider, 2003:244), the time at which the pioneer column moved to settle in Zimbabwe and set up 'military forts', a development that Schneider (2003:244) notes as illustrative of this phase. Forts that were established in Zimbabwe beginning that time include Fort Tuli, Fort Victoria, Fort Charter and Fort Salisbury. However it should be pointed out that the new settlers largely comprised South Africans of British origin who most probably had already experienced the foundation phase (and therefore the 'focussing' that Schneider (2003:244) refers to) whilst still in South Africa. On the whole, it should be pointed out that, until very recently, "the direct influence of native-speaker English on Black African English is difficult to assess" (Schmied, 1996:304) owing largely to the much smaller number of the native speakers as compared to South Africa.

According to Raftopoulos and Mlambo (2009), in 1953, Britain combined Southern Rhodesia, Northern Rhodesia and Nyasaland in a federation that was to last 10 years up to 1963. The years between 1895 and 1980 (when Zimbabwe gained its independence) may be understood as constituting the exonormative stage for ZimE, a stage when "colonies or settlers' communities tend to stabilise politically, normally under foreign, mostly British, dominance, whatever the political status may be" (Schneider, 2003:245). In the context of Zimbabwe, this was also a time that witnessed a rapid expansion in the white population (Schmied, 1996; Mlambo, 2009). As mentioned earlier, this is a period when linguistic norms are almost totally based on external norms from the homeland. However, the exonormative phase is the time when "adjustments to the local environment start to creep in" (Schneider, 2003:245). Though White Zimbabwean English is different from ZimE as used by Black second language users, it would seem that the characteristics of White Zimbabwean English observed by Fitzmaurice (2010) emanate from these stages. These include Afrikaner influences in pronunciation of consonants, raised and centralised short vowels, monophthongized diphthongs, whining voice quality (particularly among women), use of temporal adverbials, lexical borrowings from Afrikaans, Shona and Ndebele, among other characteristics (Fitzmaurice, 2010:275-283). However, it would be

anticipated that these features would also permeate ZimE as used by Black second language users since it would be the Zimbabwean whites' English that was in constant contact with that of Black Zimbabweans.

As decolonisation set in in the early 1960s (Mlambo & Raftopoulos, 2009), Britain widened participation of black majority rule, to the ire of the white settlers who voted in a 1964 referendum for independence from Britain resulting in the Unilateral Declaration of Independence (UDI) in 1965. Though Event X, "a quasi-catastrophic political event which ultimately causes the identity alignment of STL-strand speakers to switch from a self-association with the former mother country however distant, to a truly independent identity, a case of 'identity revision' triggered by the insight that one's traditional identity turns out to be 'manifestly untrue' or at least 'consistently unrewarding'" (Schneider, 2003:250) is usually a feature of the endonormative stage, UDI could be understood as some form of Event X, in the development of White Zimbabwean English.

Local activism, strikes and riots as from the 1940s culminated in a full-blown civil war in 1972 pitting against one another the Rhodesian regime army and the national liberation armies (ZANLA and ZIPRA) which were sponsored by the Zimbabwe African National Union (ZANU) and the Zimbabwe African Patriotic Union (ZAPU) political parties (Raftopoulos & Mlambo, 2009). Talks to end the war started in 1978 and the war was brought to an end in 1979 through a negotiated settlement. The British government brokered new negotiations in 1979 after which the country went for general elections in 1980, which were won by President Robert Mugabe of ZANU. In terms of Schneider's dynamic model, the attainment of independence and the dawn of majority rule in Zimbabwe in 1980 would be the true Event X if it were followed by development of an endonormative variety of English. To establish the extent to which the English was endonormative during this period would be difficult given that there were very few studies on the linguistic nature of the English of Zimbabwean Black users of English. The findings of some of these studies are detailed in 2.10 below. Studies by Magura (1984; 1985), Makoni (1993), McGinley (1987) and Ngara (1982), are among the few that attempted to address the question of the sociolinguistic status of ZimE, with the findings from the studies painting a somewhat inconsistent picture.

After a few years of political and economic stability in independent Zimbabwe, Mugabe's popularity began to decline, culminating in the rejection of a draft constitution in which an extensive programme for land reform had been proposed (Raftopoulos & Mlambo, 2009). The 2000 general election results also showed dwindling support for ZANU-PF and President Robert Mugabe. Following the rejection of the draft constitution and the narrow win by ZANU-PF in the 2000 elections, with Mugabe's blessing, the war veterans spearheaded the fast-track land reform programme during which farmland owned by about 4000 white farmers (out of 4500 farmers) was transferred to new black farmers, severely constraining relations between Zimbabwe and Britain and her Western allies (Raftopoulos & Mlambo, 2009). Relations between Zimbabwe and the World Bank and the International Monetary Fund were also affected. Most of the irked farmers left the country, considerably depleting the number of STL strand speakers, a situation which, coupled with a new sense of black empowerment would have augured well for the consolidation of a national variety of English or even a dislodging of English from its position of privilege. (As I explain below, this did not happen).

From then on, Zimbabwe's economy took a downward spiral against the backdrop of reduced agricultural production. In the 2008 harmonised elections, Morgan Tsvangirai of the Movement for Democratic Change (MDC) who narrowly beat President Mugabe pulled out of the presidential election re-run, alleging an orgy of violence against his supporters (Raftopoulos & Mlambo, 2009). In the ensuing crisis, Thabo Mbeki, then President of South Africa brokered talks between ZANU-PF and MDC-T, leading to the formation of a Government of National Unity (GNU). The term of the GNU ended with the 2013 harmonised elections which ZANU-PF won resoundingly despite allegations of rigging, a development that has seen a relapse of the economy again.

Despite this long and chequered history, the fortunes of the English language in Zimbabwe have remained favourable. The colonial and postcolonial period has seen the language being entrenched in functional domains including government administration, law, media, education, business and science and technology. Thus as Kadenge and Nkomo (2011), Mlambo (2009) and Schmied (1996) among others, also observe, Zimbabwe, along with Zambia and Malawi, is a typical English as a Second Language (ESL) country with English firmly rooted. However, some scholars, like Makoni (1993), argue that the ESL label applies only to urban parts of the country

arguing that most rural Zimbabwe is actually EFL, sociolinguistically speaking. Using Mollin's (2007) criteria for determining variety status, on the score of functional expansion alone, ZimE would pass the variety status test as a new English, a point also made in Marungudzi (2016a:274).

Government attitudes towards the English language in Zimbabwe have been consistently positive. For example, in spite of recommendations by academics such as Chimhundu (1997), Magwa (2010a; 2010b), Ngara (1982) and Nyika (2008) as well as other language promotion groups and activists for the major indigenous languages to be used as media of instruction in education, these recommendations have not seen the light of day. The print media and electronic media are also still dominated by English (Makanda, 2009; Mlambo, 2009). Entry into higher education institutions and formal employment still require a pass in English. A GCE Ordinary Level ('O' Level) certificate without English is deemed incomplete. However, the 2013 20th amendment to the constitution pronounced that all Zimbabwean indigenous languages are now officially recognised languages in order to bring them on a par with English and in September 2016, the government introduced news bulletins in previously marginalised indigenous languages on the national television broadcaster, ZBCTV.

The absence of a clear and comprehensive language policy in Zimbabwe to date (36 years after independence) is in fact a sign that language issues, deliberately or inadvertently, are not taken seriously. From a different perspective, it would appear that the Zimbabwean government is entangled in a language policy web difficult for it to come out of. In 1996, Schmied wrote that active language policy was an exception in Central Africa i.e. "In many ways, Zimbabwe, Zambia and Malawi are typical 'anglophone' countries in the sense that no coherent overall language policy was propagated (paying lip-service to African languages often went parallel with maintaining old requirements concerning English)" (Schmied, 1996:305); and for many years after independence "it was very difficult to find any comprehensive statement on the governments' educational language policies, for instance" (Schmied, 1996:305). This state of affairs has hitherto remained unchanged. For example, there is the frequent newspaper headline that decries the neglect of indigenous languages in the country. It is in this light that one is compelled to agree with Kadenge and Mudzingwa's (2011:144) prognostication that "current trends in language use in Zimbabwe suggest that the relationship between English and

indigenous languages will continue to be asymmetrical for many years to come”. As observed elsewhere such an asymmetrical relationship is fertile ground for the development of nativised varieties of English.

Schmied (1996:306) also notes that the central reason why the independent governments of the former frontline states did not abolish the colonial heritage of English relates to the then apartheid education in South Africa where public perception equated African languages to inequality and underdevelopment; and English to freedom and internationalisation. Schmied (1996:306-307) had hoped then that the end of the apartheid era would result in a reorientation of such a perception, but so far his hope has not been realised though a semblance of respect seems to be accorded to some former ‘minority’ indigenous languages (specifically Kalanga, Nambya, Ndau, Tonga, Tsonga, Sign Language, Sotho, Xhosa and Venda,) whose study and teaching has been pioneered by the Great Zimbabwe University (where this researcher is currently teaching).

As Kadenge (2010:36) argues, the traditional argument that English is a colonial relic that should be ignored is no longer tenable given that English is the only official language that Zimbabweans in particular and the rest of Africans in general can effectively utilise when communicating with other people at international fora. Arguing that English should not be seen as a threat to the Zimbabwean indigenous languages, Kadenge and Mudzingwa (2011) however remark that the language policy of Zimbabwe should foster the Zimbabwean variety of English but laments scanty research on the variety. It is, in part against this background of scanty research on the identity and status of ZimE that the present study has been conceived.

2.10 Some studies of English in Zimbabwe

Although the body of research on English in Zimbabwe is small (Kamwangamalu & Moyo, 2003; Marungudzi, 2016b; Schmied, 1996,) when compared to research on West and East Africa, some studies have been carried out to date. It is useful to classify the studies in terms of their themes and approaches. For example, there are studies that examine the history and nature of English in Zimbabwe as well as those that sought to explore the internal types of English that constitute ZimE from a sociolinguistic perspective. These two approaches would thus fit the

English studies as well as the sociolinguistic / the sociology of language approach as described in 2.4 above. There are also studies that have looked at the role of English from a language planning point of view and these can be said to be sociolinguistic in orientation as well.

There are also studies that looked at the structural characteristics of ZimE and those that looked at how Zimbabweans perceive English. Such studies have combined the English linguistics and linguistic approaches discussed in 2.4 above. These different approaches and the specific studies that have adopted them are described below.

Among those studies that adopt a sociolinguistic approach to the study of ZimE are Chiwome and Thodhlana (1992) Magura (1985), Makoni (1993), Mlambo (2009), Ngara (1977; 1982) and Schmied (1996).

Ngara (1977; 1982) chronicles the circumstances surrounding the coming of the English language to Zimbabwe and how it interacted with Shona, one of the major languages spoken in Zimbabwe. Ngara's studies seem to suggest that the language policies of the colonial settler regime are largely responsible for the current privileged status of the English language in Zimbabwe. Though scholars like Hungwe (2007) suggest that the Zimbabwean populace actually demanded English, Ngara (1982) argues that these policies imposed English as the language of education, administration, parliamentary debates and the law. In terms of the form of English used in Zimbabwe, Ngara (1982) is of the view that there is clearly an identifiable variety of English spoken in Zimbabwe, particularly at the phonological level. That new Englishes are phonologically distinct from Standard English is certainly plausible and Kadenge (2009; 2010), as discussed below, testifies to that fact.

From a related perspective but taking ZimE as part of a broader variety of Southern African Black English, Magura (1985) seeks to find out the form of English used in Zimbabwe taking the statement: "Africans are creating out of English a language of their own: a language in actions using words that dart back and forth on quick-moving feet, virile earthy and garrulous" from Themba (1959) as a cue to research on the form of Southern African Black English. Magura finds the varieties of English spoken in Zimbabwe as fitting into Platt *et al.*'s (1984) lectal range, i.e. acrolect, mesolect and basilect, and goes on to identify the groups that speak these different forms of English. Like Ngara (1982), Magura (1985) also uses sociolinguistic introspection and

anecdotal evidence for his conclusions, which as pointed out earlier, though they would give pointers to the identity of ZimE, are arguably methods that would not conclusively settle the variety status question.

Another scholar who adopts a sociolinguistic approach to the study of English in Zimbabwe is Makoni (1993). Makoni (1993) is the first researcher to address the question ‘Is Zimbabwean English a new variety of English’ (Kadenge, 2010), a question whose answer has remained debatable to date. Makoni (1993) concludes, using the criteria of development through an educational system, use of language variety for a wide range of purposes and exhibition of localised features as proposed by Platt and Weber (1980) in section 2.6 above, that the variety of English spoken in Zimbabwe is not a new type of English. Makoni (1993) props this verdict with reference to the fact that the norms of accuracy and appropriacy guiding ZimE are premised on native speaker norms, which is not the case in other types of new Englishes such as Nigerian English or Ghanaian English. Kadenge (2010: 38) takes exception to this conclusion arguing that:

...to describe Zimbabwean English as an inter-language is a purist position that has been consistently rejected by the pragmatist school of thought or the Kachruvian paradigm of world Englishes which argues that the varieties of English in the outer circle communities (Zimbabwe included) constitute different Englishes in their own right that express independent sociocultural identities and whose legitimacy owes no allegiance to the so-called native speaker norms.

Such a lack of consensus albeit based on personal introspection and conjecture betrays a need for current research to come up with more reliable methods of settling the variety status question to complement the earlier approaches.

In 2009, the question of whether ZimE is a new English was taken up by Mlambo (2009). In this study, Mlambo (2009) addresses questions that include: Is there a Zimbabwean variety of English?, If so who speaks it?, Is English in Zimbabwe an interlanguage?, Are there many varieties of English in Zimbabwe which are pragmatically identifiable as Zimbabwean?, Does the majority of Zimbabweans appear to speak an English which reflects the linguistic characteristics of Shona?. These questions are reminiscent of those posed by de Klerk (1999) discussed above about Black South African English. Though Mlambo (2009:18) does not use corpus methods to find answers to these questions, he concludes (like Magura, 1985) that “it is a

fact that there are many varieties of English in Zimbabwe which are pragmatically identifiable in Zimbabwe” i.e. native variety, near-native variety, acrolect, mesolect and basilect.

Studies that focus on English in Zimbabwe from a language planning perspective include those by Chivhanga and Chimhenga (2013), Hungwe (2007), Kadenge and Mugari (2015), Kadenge and Nkomo (2011), Magwa (2010a; 2010b), Makanda (2009), Makoni *et al.* (2008), Maseko and Dhlamini (2014), Ndamba (2013), Nhongo (2013), Nkwe and Marungudzi (2015) and Nkomo (2008). To various degrees all these studies seem to agree that there is too much attention being given to English in Zimbabwe, with the undesirable effect that both major and so-called minority indigenous languages are put in a disadvantageous position. It is this dominance of the English language in Zimbabwe that raises the question of the linguistic characteristics of English used in the country. Hasn't the English been 'domesticated', as it were, to reflect the linguistic and cultural traits of the Zimbabwean indigenous languages? If so, in what ways and with what consequences on the relationship between ZimE and those so-called Standard Englishes? In addition, though these studies do not give attention to the linguistic features and variety status of English in Zimbabwe, they provide useful background for resolving the variety status question by exploring the functions of English in Zimbabwe and attitudes towards the language.

Among the studies that focus more on the linguistic features of the English spoken in Zimbabwe are those by Fitzmaurice (2010), Kadenge (2009), Kadenge (2010), and Kadenge and Mudzingwa (2011).

Though largely focusing on the historical sociolinguistic account of the emergence and fossilization of L1 Rhodesian English (the L1 variety of a small white minority community that used to reside in Rhodesia [now Zimbabwe] that no longer exists as a coherent fragment), a study by Fitzmaurice (2010) indicates that Rhodesian English is different from ZimE and is characterised by such linguistic features as short vowels, monophthongization of long vowels and diphthongs, absence of linking 'r', 'whining' voice quality, and a lexicon comprising largely British English garnished with items from Afrikaans, Ndebele and Shona among other linguistic features.

Though this study does not particularly focus on the grammatical features of White Zimbabwean English, it is of interest in as far as it provides an STL strand of the variety that holds potential

points of comparison and contrast with the IDG strand i.e. ZimE as used by Black second language users. It will also be interesting to fathom the effect of the mass exodus of white settlers from Zimbabwe following the politico-economic crisis emerging at the turn of the millennium.

Kadenge's (2009; 2010) and Kadenge and Mudzingwa's (2011) studies largely focus on the phonological features of the English spoken in Zimbabwe and conclude that there are elements of nativisation at this level. Though Kadenge's (2009) conclusions are based on a corpus, it is important to point out that the nature of the corpus would be exclusively appropriate within the phonological tradition. Thus the present study seeks to probe beyond the phonological level and using a broader corpus modelled after the ICE framework to determine the grammatical features of the English used in Zimbabwe.

It has emerged from this section that the few research studies that have been carried on ZimE or English in Zimbabwe have indeed made attempts to engage the question regarding the ontological status and to some extent the structural features of the new English. However, there are contradictory findings emerging from these studies. The contradictory results may be attributed to the different methodologies and perspectives taken in the studies, but they point to the need for a more systematic method of settling the question. The studies have also shown a tendency to focus primarily on the phonological level of the variety to the exclusion of the morphosyntactic level. Other studies have just focussed on the role of English from a language planning perspective. Against this background, it is my conviction that an exploration of the characteristic morphosyntactic features of the ZimE using the corpus linguistics method would go a long way in clearing the air before the question of ontological status of ZimE can be settled.

2.11 Conclusion

This chapter has attempted to contextualise the study of the English used by Black second language users of English in Zimbabwe by putting it in the broader perspective of theoretical approaches to the study of world Englishes. It has discussed the relationship between new Englishes and the related phenomena of pidgins and creoles and given an overview of research on new Englishes in different parts of Africa. In terms of the sociolinguistic history and status of

English in Zimbabwe, it emerged that despite English being widely used in Zimbabwe for official purposes and communication across language groups, the nature of the English that is used has not been adequately explored to ascertain its variety status. This is in contrast to neighbouring South Africa where extensive research on Black South African English has been done. The literature reviewed in this chapter has raised pertinent issues that the present research will engage. The literature also clearly points to the existence of a research gap in terms of our knowledge about the nature and variety status of English as used by Black second language users in Zimbabwe. The present study hopes to reduce that gap.

Chapter 3

Research methodology

3.1 Introduction

As indicated in Chapter 1, the current study employs corpus linguistic methods to address the question of the linguistic features and variety status of Zimbabwean English. This chapter explores the origin, history and development of the corpus linguistic approach as well as its tenets, strengths and weaknesses. In specific terms, the chapter describes the corpus-based approach in terms of the type of data it deals with as well as the procedures and basic considerations to be born in mind when designing and building a corpus. The chapter will also describe more recent applications and influence of the 21st century technological advancements on the corpus approach. This background is intended to bring out the relevance of the corpus-based approach for the current study, particularly in contrast to largely intuition-based methods used by previous researchers. The chapter goes on to provide a description and empirical application of the methodological literature to the current study and this allows for a more nuanced evaluation of the research findings in Chapter 4.

3.2 Historical origins of corpus collection

McCarthy and O’Keeffe (2010:3-4) trace the activity of corpus collection and analysis to the 13th century when biblical scholars made attempts to compile concordances of the Bible. The aim of concordancing was to specify the words contained in the Bible for other biblical scholars. The concordances gave the citations and pages of particular chapters and verses. McCarthy and O’Keeffe (2010:3) postulate that one Anthony of Padua (1195-1231) is associated with the first known Bible concordance. The concordance by Anthony of Padua was followed by a word index of the *Vulgate* (5th C Latin version of the Bible) which was compiled by Hugo of St Caro in 1230. From there on, there came numerous Bible concordances, among them *A complete concordance to the Holy Scriptures* by Cruden in 1737 and the *Exhaustive concordance of the Bible* by Strong in 1890 (McCarthy & O’Keeffe 2010:3). Apart from the Bible, William

Shakespeare's works also became a subject of concordancing. One example of such a concordance is *A Concordance of Shakespeare* which was prepared by Beckett in 1787.

3.3 The influence of technology on corpus linguistics

According to McCarthy and O'Keeffe (2010:5), it was the revolution in hardware in the 1980s and 1990s which led to the emergence of corpus linguistics as we know it today. It is this revolution that saw the large-scale computing being done today using desktops as opposed to mainframes. The growth of the internet and fast download speeds enabled transfer of data from scholar to scholar. This also led to the availability of language data that was already in electronic format. The analogue tape recorders of the 1970s were replaced by digital recorders in the 1980s and later by high-powered video and DVD recorders.

Notwithstanding the progress in technology, efforts by linguists to harness computational power were hampered by the limitations of machines. For example, John Sinclair had to use the punched-card system (an 18th C method) in his earliest exploratory years towards the COBUILD project. Equally limiting was also the Oxford Concordance Programme of the 1980s and 1990s which tolerated no errors (even of the slightest nature) in the writing of commands. Such limitations were obviated by more user-friendly software such as the GUI-based software on which Scott's *Wordsmith Tools* (1996 -) were based. Such software is powerful, easy to use and more than up to the tasks that researchers demand of them (McCarthy & O'Keeffe, 2010: 6).

Tognini Bonelli (2010) summarises the contribution of technology to the growth of corpus linguistics from the 1960s to the present and divide this influence into three generations namely:

- The first 20 years i.e. 1960-1980 (characterised by learning how to build and maintain one-million word corpora and in manual form)
- The second twenty years i.e. 1980-2000 (characterised by 20 million word corpora targets with the help of the scanner , and then later, the First Serendipity when corpora of beyond 20 million words could be compiled courtesy of computer type-setting)

- The new millennium i.e. 2000-present (characterised by the Second Serendipity when there was a proliferation of text that never existed in hard copy form leading to compilation of corpora of unlimited quantities thanks to the internet)

3.4 Motivation for modern corpus collection

One characteristic that unites traditional and modern corpus collection is the motivation for collecting the corpus. Just as Biblical scholars collected corpora to build indexes and concordances, modern corpus collection is also motivated by the need to create databases of processed texts for different types of analysis. Modern corpus collection has its roots in the work of Roberto Bus, a Jesuit priest who created an electronic index of St Thomas Aquinas between 1950 and 1970 (McCarthy & O’Keeffe, 2010:4). Corpus collection was also influenced by the work of lexicographers as well as by the work of pre-Chomskyan structural linguists. In the early years of corpus collection, corpora were compiled on paper. Samuel Johnson’s first comprehensive dictionary was compiled using a paper corpus. A paper corpus was also used to compile the *Oxford English Dictionary*. McCarthy and O’Keeffe (2010:4) report that the corpus for the *Oxford English Dictionary* amassed more than 3 million slips of paper attesting word usage. However collection of real data came of age in the 1950s in the era of American structuralists such as Harris, Fries and Hill. The main motivation for corpus collection by structuralists was “to put real language data at the core of what linguists study” (McCarthy & O’Keeffe, 2010:4). Thus structuralists were the forerunners of corpus linguistics. So “where the work of the early biblical and literary scholars provides the background *modus operandi* of word searching and indexing, the structuralists were the forerunners of corpora” (McCarthy & O’Keeffe, 2010:4) not only in the sense of data gathering but in terms of the commitment to putting real language data at the core of what linguists study. One instance of language study based on real language data was the growing of interest in language acquisition based on transcribed data, which dates back to the 1960s. This resulted in the CHILDES Language Database.

McCarthy and O’Keeffe (2010:4) go on to report that the first computer-generated concordances appeared in the late 1950s and used punched-card technology for storage. By 1979, computer-

generated concordances were already seen as a “general-purpose working tool for the study of literature” (Howard-Hill in McCarthy & O’Keeffe, 2010:4). It was the 1980s and 1990s which really saw the arrival of corpora as tools for the linguist or applied linguist as we know them today. Prior to this linguistic terminology the word ‘corpus’ had been used for a long time to refer to a collection or binding together of written works of a similar nature.

In linguistic literature “corpus was first used by W.S. Allen in *Transactions of the Philological Society* in 1956 and was used to refer to the familiar ‘body of written or spoken material’ upon which a linguistic analysis is based” (*Oxford English Dictionary* 2009). McCarthy and O’Keeffe (2010:5) note that ‘corpus linguistics’ came into common usage in the early 1980s and its wider popularisation is attributed to Aarts and Meijs (1984).

3.5 Definition of corpora and types of corpora

So far, the words ‘corpus’ and ‘corpora’ have been mentioned a number of times in this thesis, mainly in the context of the history and justification for using corpora. Though the above paragraph points out that a corpus is a body of written or spoken material collected for the purposes of linguistic analysis, it would be important to give a more systematic characterisation of corpora. According to Cheng (2012:3), a corpus “is a collection of texts based on a set of design criteria one of which is that the corpus aims to be representative”. So a corpus is not a collection of texts regardless of type of texts and weightings attached to each text type (Cheng, 2012:3). It is also important to add that “while the term *corpus* merely refers to a body of electronically encoded text, it is not the case that a corpus consists of any collection of texts, picked at random” (Baker, 2006:26). Biber *et al.* (1998:246) make a similar point and add that “a corpus seeks to represent a language or some part of a language” and that the design for a corpus should depend on what it is intended to represent.

Sinclair (2004:16) defines a corpus as an electronic collection of language texts. Sinclair (2004:16) adds that these texts are selected according to external criteria to represent, as far as possible, a language or language variety and are used as a source of data for linguistic research. It may be remarked here that Sinclair (2004) espouses a technology-influenced definition, but as

we have seen above, in the past corpora were not always electronic. Otherwise what emerges from the definitions outlined above is the centrality of being principled and astute when designing and building a corpus. To provide further clarity on what a corpus exactly is, it is important to distinguish it from archives or databases as well as other types of text collections. Section 3.5.1 focusses on these distinctions while section 3.5.2 focusses on corpus typology

3.5.1 Corpora, archives and other databases

Baker (2006:26) uses the concept of ‘representativeness’ to distinguish a corpus from an archive or database. He points out that a corpus is designed for a particular ‘representative function’ while on the other hand, an archive or database is simply a text repository, often huge and opportunistically collected and normally not structured (Baker, 2006:26). Thus corpora tend towards having a more balanced, carefully thought out collection of texts that are representative of a language variety or genre. In reality some corpora may be collected with a degree of opportunism while on the other hand there could be some form of selection and care in constructing an archive, thus theoretically, the distinction between a corpus and a database is sometimes blurred. However, in corpus building, there should be an effort to keep in mind the idea of building a standard reference on the part of the corpus builder.

Against the background that corpora should be representative of a language or language variety, sampling is clearly an important factor in corpus building. Sinclair (2004:1) is of the opinion that despite the unsteadiness of the notion of ‘representativeness’, it remains an unavoidable notion in corpus design, along with ‘sample’ and ‘balance’. Thus care was taken in the present study to ensure representativeness and balance as will be elaborated in section 3.11.

Sinclair (2004:15) also distinguishes a corpus from related databases such as the *World Wide Web*, as well as from collections of citations or collections of quotations, and from a text. The major reasons Sinclair (2004:15) gives for not taking these as corpora are that they are not constructed from a linguistic point of view; they are not constructed on external criteria and their purpose does not focus particularly on linguistic analysis, considerations all too important in building corpora.

3.5.2 Types of corpora

Scholars categorise and name corpora following different parameters. From a broad perspective, Lee (2010:109) classifies corpora into general language corpora and specialised language corpora. Under general language corpora, he includes speech corpora, parsed corpora and historical corpora. On specialised corpora, Lee (2010:114) says that these are corpora “that do not aim to comprehensively represent a language as a whole, but only specialised segments of it” and goes on to give the Michigan Corpus of Academic Spoken English (MICASE) and the British Academic Spoken English (BASE) as examples of this type. Lee (2010:114-118) also gives other types of corpora and their examples, including multimedia corpora, developmental, learner and lingua franca corpora. It is evident from Lee’s (2010) categorisation that it is based on the inherent characteristics of the corpus (as opposed to its functions).

Tognini Bonelli (2010:22-27) admits that the proliferation of corpora now makes it difficult to classify them in any conclusive way. Thus she sketchily classifies corpora into sample corpora, corpora for comparison, special corpora, corpora along the time dimension, bilingual and multilingual corpora, non-native speaker corpora and spoken corpora. It is clear that Tognini Bonelli (2010) shares Lee’s (2010) conceptualisation that the classification of corpora is based on their indexical characteristics.

From a different perspective, a typology proposed by Baker (2006:26-30) encompasses four categories that include specialised corpora, diachronic corpora, reference corpora and monitor corpora. Baker (2006:26) understands a specialised corpus as a corpus that conforms to a specialised criterion e.g. academic essays or newspaper reports and in concurrence with Lee (2010), he observes that a specialised corpus is used to study aspects of a particular variety or genre of language. Restrictions in terms of place and time may also be put on this type of corpus. A diachronic corpus on the other hand is a corpus which has been built to be representative of a language or language variety over a particular period of time, enabling the tracking of linguistic changes (Baker, 2006:29). The ARCHER corpus and the Brown family of corpora are examples of diachronic corpora. On the other hand, a reference corpus is a large corpus consisting of millions of words from a wide range of texts e.g. British National Corpus (BNC), the American National Corpus, and the International Corpus of English. A reference corpus “acts as a good

benchmark of what is ‘normal’ in language, by which one’s own data can be compared to” (Baker, 2006:43). Finally a monitor corpus is one to which texts are continually added e.g. the Bank of English (BoE) corpus, the Corpus of Contemporary American English (COCA) and the Global Web-based English (GloWbE). It is clear that Baker’s typology is based largely on the function to which the corpora are put.

3.6 Reasons for building one’s own corpus

In spite of the existence of many corpora that may be used for reference purposes, Baker (2006:25), Reppen (2010:31) and Nelson (2010:53) point out that there are still a number of reasons why individual researchers may need to build their own corpora. Firstly, the existing reference corpora may not contain enough of the text types the researcher wants. Secondly, reference corpora may not have enough reference to the subject being investigated and may actually contain data that is already too old (Reppen, 2010:31). Baker (2006:25) is of the opinion that some of the potential problems in dealing with corpora which are by their very nature instances of decontextualised language, is a good reason for the researcher (particularly in discourse analysis) to participate in the building of his/her corpus.

Against this background, it is clear that building one’s own corpus affords the researcher the opportunity to familiarise himself or herself with the data through the researcher engaging in finding, choosing, checking the corpora, annotating the data and many other processes. In addition to that, initial hypotheses about the data may also be formulated at the stage of corpus collection, something that would be impossible if the researcher were using an existing corpus. According to Baker (2006:25), “the process of finding and selecting texts, obtaining permissions, transferring to electronic format, checking and annotating files will result in the researcher gaining a much better ‘feel’ for the data and its idiosyncrasies”, so that the researcher will not start “from the position of *tabula rasa*,” as it were (Baker, 2006:25).

3.7 Basic considerations to be made when building a corpus

The first step towards constructing a corpus is to have a clearly articulated question (Reppen, 2010: 31). Once the question is decided upon, the following step is to ensure that there is not an existing corpus that meets the needs of the research question. This measure precludes the researcher from reduplicating a corpus that is already in existence, thus saves the researcher valuable time and effort. It is also important to ensure that the corpus is representative of the language being studied (Reppen, 2010:31), a point emphasised in section 3.5.1 above. For example, if the researcher intends to study newspaper language, then the corpus should adequately represent that particular genre.

Another important consideration to be made when building a corpus has to do with the corpus size. There is no simple mechanism to determine the corpus size as this depends on the purpose of the research. According to Biber *et al.* (1998:247-248), the question of corpus size does not only focus on number of words in the corpus but also “on the number of texts from different categories, the number of samples from each text, and the number of words in each sample.” In terms of total number of words in a corpus, a one-million word corpus of one variety of language (for non-discourse purposes) is acceptable for comparison purposes (Baker, 2006:27-28). Other suggestions are that a million words is adequate for grammatical analysis, while half a million words are adequate for verb-form morphology and 100 000 words are adequate for prosody studies (Kennedy, 1998:68). On the other hand, 1 million words are too few for lexicographical purposes. For discourse analysis, corpora may be considerably short e.g. 300 – 500 words or even fewer (Baker, 2006:28). Corpus size thus very much “depends on the type of language that is being investigated” (Baker, 2006:31). Baker (2006:31) also observes that if the use of language being investigated is very specific, then there is little need to collect millions of words of data. On the other hand, if one intends to study language use in a relatively general context, it is more sensible to make use of an existing reference corpus, rather than undertake the time-consuming task of creating one’s own corpus.

Furthermore, it is always important to remember that when dealing with the corpus size question, the one-size-fits-all approach would be irrelevant (Reppen, 2010:32). If the corpus is being built for the purposes of compiling a dictionary, tens or hundreds of millions of words may be

required (Reppen, 2010:32). According to Reppen (2010:32), the size of the corpus is also determined by two important factors, namely, representativeness and practicality (issues such as time, financial and resources constraints). On the question of representativeness, the corpus may represent the language being investigated totally or partially. Section 3.10 below takes up the issue of corpus size in more detail.

In terms of the number of texts from different categories, there is need to ensure that a sufficient number of texts is included in the corpus. If the texts are too few, “a single text can have an undue influence on the results of an analysis” (Biber *et al.*, 1998:249), leading to unreliability of findings. The same astuteness is required when dealing with the question of the number of samples from each text type and that of the number of words in each sample.

According to Adolphs and Carter (2013:17-18), the following procedures summarise the practical programme that has to be followed when setting out to collect a corpus:

1. Articulate a research question
2. Identify the texts and develop the plan for collection
3. Get consent from parties concerned
4. Check rules regarding consent since these vary by country, institution or setting
5. No need for permission for data in the public domain
6. Respect copyright laws even when using material for free
7. Ask questions such as: What constitutes a text? How will the files be named? What information will be included in each file? How will the texts be stored (file format) before setting out to collect the corpus
8. Collect files at the smallest ‘unit’ since it is easier to combine files when analysing.
9. Establish file-naming conventions before saving the files, relating the name to the content of the file
10. Use headers with a consistent format (that contain information about the file) in all files in the corpus
11. Save files in plain text
12. Preserve non-standard spelling in novice writing
13. Decide what to do with non-orthographic markings or art

14. Decide how to transcribe reduced forms, inaudible parts, overlapping speech as well as conversation facilitators such as uh, huh, mmm, hum etc. and laughter.

From a related perspective, Baker (2006:31) also suggests a number of issues that need to be considered when building a corpus from scratch. These include the need to carry out a pilot study (this is necessary in order to establish what texts are available, how easily accessible and easily convertible to electronic format they are), the need to prefer using data that is already in electronic format and the need to ‘clean’ texts downloaded from the internet (removing unwanted text and images) so that they can be processed by concordance programmes such as *Wordsmith Tools*.

In cases where paper texts have been selected for inclusion into the corpus, Baker (2006:31) notes that there will be a need to convert them into electronic documents. The conversion is usually accomplished through scanning or keying the paper texts into the computer. Scanning involves running texts through a scanner with Optical Character Recognition (OCR) software to convert them from paper documents to electronic ones. If the paper texts are scanned, the scanned texts need manual checking, spell checking and correction for errors.

In addition to these guidelines, there are also issues that pertain specifically to the construction of spoken corpora. These include recording, transcribing, coding and mark-up, and management and analysis. Corpus collection also needs to be continually reassessed though this should not be construed as discounting the planning that would have been done prior to the inception of corpus collection. Adolphs and Carter (2013:8) propose the use of a checklist or log to chart progress since it acts as “an invaluable point of reference for discussing anomalies or absences that may occur in the data”. Above all, it is important to guarantee that the corpus construction process is as systematic, principled and replicable as possible.

3.7.1 Recording

Baker (2006), Reppen (2010) and Adolphs and Carter (2013) spell out a number of requirements when it comes to recording oral texts. The authors concur that recordings should be of good quality, sufficiently rich and varied. Metadata also needs to be added to the recordings (editorial

metadata, analytic metadata, descriptive metadata and administrative metadata) so that representativeness, balance and homogeneity may be achieved. Ethical issues of consent and anonymization are also critical. Adolphs and Carter (2013: 10-11) specifically warn against alteration of word output for the purpose of anonymization arguing that this “can make for an inauthentic record and render the data unsuitable for naturalistic and phonetic and prosodic analysis”. A similar problem arises with the use of video data.

3.7.2 Transcribing

When transcribing, researchers need to decide on the amount of detail to be included in the written record. There are a number of transcription conventions e.g. Text Encoding Initiative (TEI) and the Network of European Reference Corpora (NERC). Key features that most users of spoken data will want to see marked up include who is speaking, when and where, and to whom; interruptions, overlaps, back channels, hesitations, pauses, and laughter, as they occur in the discourse, as well as some distinct pronunciation and prosodic variations (Adolphs & Carter, 2013:12-13). There are three ways of transcribing recordings, namely, linear representation, column-based transcript and musical score-type format. It would also be desirable to have software programmes that are capable of transcribing multimodal data but, as Adolphs and Carter (2013:15) point out

the alignment of the different elements and the software needed to analyze such a multimodal resource are still in the early stages of development, and at the present time it is probably beyond the scope of the majority of individual corpus projects to develop a searchable resource that includes the kind of dynamic representation that would address the need for a less linear transcription layout.

3.7.3 Coding

Coding refers to “the assignment of events to stipulated symbolic categories” (Adolphs & Carter 2013:15, citing Bird & Lieberman 2001). The authors add that coding is essentially a development of the transcription stage, providing further detail to supplement the basic systems of annotation and mark-up that are applied through the use of transcription notation. The coding

stage thus operates at a higher level of abstraction compared to the transcription stage and may include, *inter alia*, annotation of grammatical, semantic, pragmatic or discursal features.

Coding involves manually or automatically marking up and tagging discourse transcripts. Discourse features such as demographic information, extralinguistic information, part-of-speech (POS) information, prosodic features, phonetic features or a combination of these may be included in the coding. Defining mark-up as the process of “adding information to a corpus file,” Reppen (2010:35) posits that mark-up subdivides into two categories i.e. document mark-up and annotations. Document mark-up refers to markings like Hyper Text Markup Language (HTML) codes which are used to indicate document features such as paragraphs, fonts, sentences, sentence numbers and speaker identification. Headers are another example of document mark-up for they provide additional information about the text. Reppen (2010:33) observes that headers contain information such as the demographic detail about the writer or the speaker, or contextual information about the text such as when and where the text was collected. Headers should always be formatted in a uniform way and comprehensive information should be given so that no detail is missed when it is time for analysis.

On the other hand, annotation involves parts of speech tagging, as well as the tagging of syntactic, semantic and prosodic information. The same corpus can be marked up with multiple types of mark-ups and the relevant mark-up types and the version with the relevant type of mark-up or combination of mark-up types may be viewed at a given time. Annotation also involves the creation of *headers* and *body* of text. Headers contain information such as the author, genre of the file, title, date of publication, place of publication, medium of text etc. (Reppen, 2010:33). Such information is useful at a later analysis stage. Annotation in the body of the text involves indicating the grammatical classes of the words in the text, verb tenses, number of nouns (singular or plural). Such grammatical annotation enables corpus users to make more specific analyses. Annotation may also involve the encoding of semantic information

Using a broader term ‘annotation’, Baker (2006:38) notes that coding is necessary in the sense that it will be easier to keep the structure of the corpus as well as to aid analysis. A standardized coding system should be used. According to Baker (2006:39), one example of such a standardised system is the Standard General Markup Language (SGML). This system uses codes

which are also known informally as *tags* and more formally as *elements*. Apart from HTML, other forms of SGML coding are Text Encoding Initiative (TEI) and eXtensible Markup Language (XML). Baker (2006:39) observes that it is important to be familiar with these coding systems so that they can be used when the need arises (not to invent one's own idiosyncratic annotation system, which is actually a shame and may not be handled by corpus analysis packages such as *WordSmith Tools*).

Nevertheless, discrimination should be exercised in deciding whether to annotate a corpus or not; or what type of annotation to employ, depending on the research questions being investigated (Baker, 2006:40). So annotation is not necessarily a compulsory process. It is also possible to come back to the corpus to annotate it in a particular way at the point where the need arises (Baker, 2006:40).

3.8 Limitations in using corpora for research purposes

In spite of the acknowledged strengths of the corpus-based approach in language research and the ways it can be used to complement the shortcomings of participant observation and other forms of anecdotal evidence as illustrated in the foregoing discussion as well as in Chapter 1, many scholars have pointed to the limitations of this approach. For example, Reppen (2010:31) has pointed out that the much treasured representativeness requirement expected of most corpora (required chiefly because of the finite size of corpora versus the open-endedness of actual language in use) is not possible to achieve in most cases. Biber *et al.* (1998:246) submit that representing a language or a part of a language is a problematic task because “we do not know the full extent of variation in languages or all the contextual variables that need to be covered in order to capture all variation in texts”.

From a related perspective, Sinclair (2004:2) has pointed out that “no corpus, no matter how large, how carefully designed, can have exactly the same characteristics as the language itself.” Echoing similar sentiments on representativeness, Nelson (2010: 60) points out that “any attempt at corpus creation is therefore a compromise between the hoped for and the achievable”. Such a compromise is likely to result in corpora coming short of the appropriate size, missing the

occurrence of particular linguistic structures in the process. A scenario like this creates problems in the generalisation of findings. To temper the negative effects of the corpus approach, Sinclair (2004:2) suggests that we should not totally give up our intuitive judgements as language researchers. In actual fact, the linguistic facts that corpus-based research seek to verify have been raised through participant observation and intuition.

Baker (2006:25) notes a number of disadvantages in using corpora as the only source of research data, among them the fact that a corpus text is decontextualised – issues to do with the production and reception of the text cannot be ascertained simply by looking at frequency lists since a corpus is a self-contained entity. Baker (2006:25) goes on to say that quantitative analyses employed with corpora obscure relationships between sentences in the same file.

This background has two implications for the implementation of corpus-based research. Firstly, as Sinclair (2004) points out, corpus-based methods should be viewed as complementary to earlier language research approaches, and not as purely stand-alone methodologies. Secondly, the findings derived through corpus-based methods have to be interpreted within the context of the nature of the corpus as well as the corpus-collection procedures so that generalisations to larger contexts of language use are well-considered.

3.9 The possibilities of corpus linguistics in non-native English contexts

Corpus linguistics (CL) has a number of applications. It may be an end in itself whereby it provides a means for the empirical analysis of language and in so doing adds to its definition and description in terms of lexis and grammar. Besides, CL may be used from an applied perspective in the pursuit of broader research questions in areas such as language teaching and learning, discourse analysis, literary stylistics, forensic linguistics, pragmatics, speech technology, sociolinguistics and health communication (McCarthy & O’Keeffe, 2010). From this perspective, CL is a means of doing things – a means to an end rather than an end in itself. Biber (2010:159) observes that there is in fact a sense in which corpus linguistics may be understood as a methodological approach and not necessarily a model of language. Biber (2010:160) adds that the methodological innovations that are a hallmark of the corpus linguistics approach “have

enabled researchers to ask fundamentally different kinds of research questions, sometimes resulting in radically different perspectives on language variation and use from those taken in previous research”.

From a sociolinguistic perspective, CL has been used in the study of linguistic variation. In the study of linguistic variation, CL has not only been used by researchers on register and style in native language settings, but has also been used as a basis for modern dictionaries and grammars. There is a belief that CL can also be applied in the compilation of dictionaries in non-native ESL settings if corpora for such purposes were available and if researchers had interest in the undertaking (Schmied, 1991:256). However, Schmied (1991:257) points out that “there are smaller and more immediately practical corpus-linguistic projects possible that may be considered by individual ESL researchers” though he concedes that transposing the principles of corpus compilation developed for native communities to non-native communities such as those in Africa poses a number of difficulties.

Schmied (1990:258) notes that among the projects in which CL may be harnessed is one focusing on the practical grammatical and lexicographic description of non-native Englishes in which the internal structure of the ESL lexicon, the syntactic behaviour of particular words, the semantic expansion or reduction of ‘standard’ English words, inter alia, may be established. Schmied (1990:258) adds that concordance programmes can be used for such purposes.

Schmied (1990:258) also notes that more sophisticated statistical programmes such as the *SPSS/PC+* and analytical tools such as *Word Cruncher* can be used in semantic-syntactical, textual and sociolinguistic analyses. Such analyses can establish the density of variety-specific features and enable comparison of different text genres in a particular language variety.

From such corpus applications, it is evident that the CL approach can make contributions to research into the phenomenon of nativisation. One enduring advantage of this approach is that it allows the researcher “to confirm quantitatively and statistically impressions he has gained from introspections or participant observation about qualitative differences between native and non-native Englishes” (Schmied, 1990:259).

If grammatical analyses are done, it would be possible to prove that “non-native Englishes form clines of intelligibility and develop their own systematicity” (Schmied, 1990:259), thus complementing the usually qualitative focus of non-native English studies with a quantitative dimension. It is also anticipated that the CL approach should be able to extract information about an unknown writer on purely formal textual criteria.

For the purposes of this study, as indicated in Chapter 1, the CL approach has been used to explore the linguistic features and the variety status of the variety of English used by black second language users in Zimbabwe. Section 3.11 details the methodological aspects of the current study.

3.10 Building a corpus to represent a variety of a language

The term variety is so broad; it encompasses national varieties, regional varieties and styles. Therefore “the starting point for the building of a variety of a language could usefully be based on a fundamental decision: is the proposed corpus being built to represent a *Variety* of a language, such as American English or British English or is to represent a *variety* such as legal English or academic English” (Clancy, 2010:81). A *Variety* is geographically defined and is user-related whereas a *variety* is defined situationally and is use-related. It is important to point out here that it is the *variety* that is also known as *register* by scholars such as Biber *et al.* (1998), Wardhaugh (1998), Mooney *et al.* (2011). As emphasised in earlier sections, of key importance to building a corpus for a *Variety* or *variety* of language is the concept of representativeness. Thus it is of critical importance to design a corpus that represents the language variety as a whole. So, issues of corpus size, diversity of texts, number and length of texts should be carefully considered though other factors outside the control of the corpus builder may affect these ideals. The attributes of the corpus that was collected for the purposes of this study are detailed in 3.11 below.

All these issues however depend on whether the corpora are used for sociolinguistic purposes or for applied linguistic purposes (Clancy, 2010). This is because the conceptualisation of language variation by sociolinguists differs from that by corpus linguists. Whereas sociolinguists

conceptualise language varieties in terms of language use variation engendered or traceable to factors such as age, social class, sex etc., corpus linguists conceptualise language variation in relation to circumstances and purpose (register variation or genre variation). Corpus linguists are also more concerned with what the speakers are doing rather than who the speakers are while on the other hand sociolinguists are concerned with who the speakers are i.e. their gender, ethnicity, age etc.

As noted above, in designing a corpus for a *variety/Variety* of a language, a number of issues should be considered, namely, corpus size, diversity of texts to include as well as length and number of texts.

In terms of corpus size, as noted earlier, there is no definite answer that may be given regarding the size of the corpus since this depends on the amount of time available as well as other resources. According to Clancy (2010: 82), the guiding principle is ‘cut the coat according to the cloth’. It is also important at this point to decide whether the corpus will consist of spoken or written texts or both. As pointed out earlier considerations of corpus size are also intricately tied with those of the number of texts from different categories, the number of samples from each text and the number of words in each sample (Biber *et al.*, 1998:249). Enough texts should be included in each category because “if too few texts are included, a single text can have an undue influence on the results of an analysis” (Biber *et al.*, 1998:249).

As for diversity of texts, a corpus whose purpose is variegated needs to contain a very wide range of texts (Clancy, 2010:84). Similarly, Biber *et al.* (1998:248) say that “any corpus that is used for studies of variation or that seeks to represent a language needs to be concerned with the diversity of texts” there being no such thing as ‘general language’. Using a diverse range of texts is vital in the sense that there are important linguistic differences across different varieties of language. Important areas of concern in issues of diversity include register variation, dialect and subject matter. There is need for variety in all these areas. Different corpora in existence i.e. the Brown Corpus, the LOB corpus, the CANCODE corpus, the ICE corpus, the BNC corpus, among others, have used different sampling techniques to achieve the required diversity.

Regarding length and number of texts, it is important to strike a balance between the length and number of texts of different genres. If the types of texts are not balanced, chances are that, as

already pointed out, the corpus will be skewed by relatively long texts. Sinclair (2005 in Clancy 2010:85) suggests that one way to prevent such skewedness, “is to build a corpus large enough to dilute even the lengthiest text”. It is also evident here that the question of text length and number cannot be divorced from the question of corpus size. It should also be important for researchers to include more different types of texts rather than lengthier text samples (Meyer in Clancy, 2010:85). According to Biber *et al* (1998:249), “for all research, it is important to realize that size cannot make up for the lack of diversity”. It is also important to balance initial, medial and final parts of texts. As for length of texts, Clancy (2010:85) suggests that 2000 – 5000 words and for number, 8 – 80 texts per category was found through research to be representative enough. Another consideration to be made is that texts should also end at suitable discourse breaks.

With reference to Leech (1991), Clancy (2010: 86) outlines six criteria that should be followed in order to achieve corpus representativeness and balance, of which the following are germane to this study:

- Structural criteria should be identified for the construction of the corpus, from where a framework of main components should be derived.
- An inventory of text types should be set out that should be included in each component.
- The size of each component, the number of texts and the size of individual samples should be determined.
- Adequate documentation should be kept, to allow future users of the corpus to interpret their results in terms of the corpus composition.

Against this background this chapter in general describes these corpus building methodological specifics and in particular section 3.11 documents the corpus design in terms of sampling, data-collection procedures, annotation and file-naming practices as well as corpus analysis techniques.

3.11 The corpus design of the present study

3.11.1 The research questions addressed by the corpus study

It should be restated that the current study intends to explore the characteristic morphosyntactic features of a variety of English i.e. ZimE as used by black second language users in Zimbabwe. In terms of the research question, it can be stated, as indicated in Chapter 1, that this study intends to explore the morphosyntactic features characteristic of this variety as well as the variety status of English in Zimbabwe. It was also determined that these questions can be best explored using oral and written texts in the English variety used in Zimbabwe. It is therefore clear that there was a need to collect a corpus of ZimE, unavoidably so because there existed, at the time of the inception of the study, no appropriate corpus on ZimE known to me. Chapter 1 has made reference to a corpus by Louw and Jordan (1993) as well as a more recent but small phonological corpus by Kadenge (2010). The corpus by Kadenge (2010) would hardly have been useful in the exploration of morphosyntactic features of ZimE, since it is constituted by decontextualised recordings of single English words/sentences intended to examine pronunciation features. The corpus by Louw and Jordan (1993) as reported in Chapter 1 also lacked the diversity and variety representativeness, as it is constituted by only two four-book series of secondary English readers. It also needs to be restated that the features of ZimE that need to be investigated are largely of a morphosyntactic nature.

3.11.2 The corpus sample

It emerged in the discussion of the corpus methodology in the above sections that the question of corpus size is not amenable to an easy answer (see sections 3.7 and 3.10). It should be stated that the corpus size used for this study was mainly determined by time and resource constraints. Various corpus linguists believe that the biggest corpus is the best but this study has been guided by more recent research studies that indicate that diversity of texts is more important than corpus size. Thus a corpus of about 390 000 words was collected. In terms of the types of corpora described in section 3.5.2 above, this corpus is something between a specialised corpus and a reference corpus. It is a reference corpus in the sense that it attempts to represent a language

Variety, though it is not comparable to existing reference corpora in terms of their large sizes (with some now running into billions of words). From another perspective, it is a specialised corpus in that it has a very specific purpose; it is not very large though it does not focus only on a particular variety of language, but on a number of registers.

Moreover, in order to ensure the representativeness mentioned in the guidelines for corpus collection discussed above, 195 texts representing various registers were selected, following the ICE conventions (to some extent) in order to represent the variety of English used by black second language users in Zimbabwe. Table 3.1 summarises the corpus sample size and the major subcategories for the different texts:

Table 3.1: The corpus sample

Corpus mode	Genre	Quantity
Spoken (103 = 53.4%)	Private dialogue	12 samples
	Public dialogue	48 samples
	Public unscripted monologue	25 samples
	Public scripted monologue	18 samples
Written (92 = 46.6%)	Non-printed student writing	15 samples
	Non-printed letters	12 samples
	Academic writing	15 samples
	Popular writing	14 samples
	Newspaper reportage	18 samples
	Creative writing	18 samples

Table 3.1 shows that the corpus consisted of both spoken and written texts. This balance is an important aspect since a corpus intended to represent a language variety should sample texts from both modes of language use. According to Sinclair (2004:8), the proportion of the spoken mode to the written mode may vary between 50 per cent and 90 per cent though most reference corpora including the ICE corpus have adopted the 60 versus 40 per cent ratio. Table 3.1 shows

that the spoken mode of the corpus constitutes 53.4 per cent while the written mode constitutes 46.6 per cent. Sinclair (2004:8) observes that the tilt towards the spoken mode of language adopted in most corpora follows a general guess that “most people experience many times as much speech as writing”, a caveat that may still need to be re-examined in ESL and EFL contexts as I argue below.

In terms of the spoken mode of the corpus, four subcategories (following the ICE conventions) of the corpus were adopted, namely spoken private dialogue, spoken public dialogue, spoken unscripted monologue and spoken scripted monologue. It should be pointed out that though an attempt was made to balance between the spoken private dialogue and the spoken public dialogue, instances where second language users of English in Zimbabwe used English in private direct conversations were largely hard to come by. (Some of the possible explanations for this are given in Chapter 4.) This explains why there are 12 texts of spoken private dialogue versus 48 texts for spoken public dialogue. Though not achieving a perfect balance, the written mode presents a slightly more balanced distribution of different genres of written language. Table 3.1 shows that there are subcategories including non-printed student writing, non-printed letters, academic writing, popular writing, newspaper reportage and creative writing. Following the principle of balance, these six genres were in turn divided into respective subcategories. For example, non-printed student writing was grouped into disciplines such as natural sciences, social sciences and human sciences, non-printed letters into social letters and business letters and so on.

In terms of diversity, as explained above, it is evident from the description that the sample represents a sufficiently wide range of the use of the variety of English used by black second language users in Zimbabwe, both in the oral and written modes as well as in the formal and informal style of the variety.

As already mentioned, it should be born in mind that the ICE corpus has used a 60-40 ratio for the spoken and the written component of the corpus. This appears to be in order in countries where speakers produce and receive more spoken language than written language. In the context of Zimbabwe, people generally experience more English in the form of writing than in the form of speech, unlike in most English-speaking countries for which English is used in speech by the

majority of the population. So the ICE model may not be very representative of the Zimbabwean context.

3.11.3 Data collection procedures

The data used for this study falls into the oral, written and visual modes. However, for the visual data, no effort was made to include the visual aspects of the texts into the analysis because, though the visual aspects might have had potential to enhance the discourse context of the texts, representing these aspects in transcription and analysis would have posed challenges (Adolphs & Carter, 2013:11-12). Before a full-scale collection of the data was embarked on, a pilot study, as recommended by Nelson (2010) and Baker (2006), among other authors, was conducted.

A number of lessons were learnt from the pilot study. For example, it was learnt that handwritten letters (both social and business) were not easy to access. Following this discovery, the researcher settled on using electronic versions of both social and business letters, which were reckoned, and actually turned out to be, more readily accessible.

It also emerged from the pilot study that there were very few private conversations in informal contexts that were carried out in English in Zimbabwe, English being a second language in the country. This created some difficulty as I ended up with rather few private conversations. These conversations also had large portions of code switching in which the speakers used the indigenous languages of Zimbabwe, notably Shona. However, from a different perspective, this state of affairs was interpreted to be more reflective of the English discourse in Zimbabwe. Ordinary conversations between members of the public in informal settings were seen to be largely carried out in the speakers' mother tongues. Schmied (1991) reports on similar challenges in compiling the East Africa component of the ICE corpus. He thus recommends that it is sometimes necessary to depart from the parameters of the ICE in order to circumvent the practical constraints encountered in the collection of non-native ESL corpora – “one solution to these partly conflicting principles in corpus compilation seems to be a compromise” (Schmied, 1991:262). Schmied (1991:262) also recommends a multidimensional approach in the categorisation of texts.

The pilot study also awakened the researcher to the existence of a number of pre-recorded oral data which the researcher needed only to sample and transcribe. In line with the suggestion by Nelson (2010:61) that “when it comes to private data, it is always useful for the researcher to use personal contacts to gain access to the data needed for corpus building” the pre-recorded oral data largely came from the researcher’s acquaintances in the journalism industry as well as personal contacts and were in the form of CDs, DVDs and computer files.

It also emerged from the pilot study that securing consent to record oral data and satisfying all ethical requirements did not pose a big challenge, as respondents were largely co-operative. The full-scale collection of the data used for the main study was collected (including transcribing and annotation) between February 2013 and December 2015. The sections below detail the procedures that were followed in the collection and analysis of the spoken and written data.

3.11.3.1 The collection of spoken data

Spoken corpora provide a unique resource for the exploration of naturally occurring discourse (Adolphs & Carter, 2013:5), which is valuable to a diverse number of research questions/communities. Adolphs and Carter (2013:6) add that spoken corpora “provide researchers with rich samples of spoken language-in-use, which form the basis of new and emerging naturally occurring discourse”. Nevertheless the collection of spoken texts creates unique challenges. For example, all spoken data needs to be transcribed, a lengthier process than keying in or scanning. Secondly parts of oral texts may be unclear and there may be problems of overlapping dialogue as well as problems dealing with additional but irrelevant information such as volume, speed, stress, (prosodic features), coughing, laughter etc. (paralinguistic information) (Baker, 2006:35).

Regarding unclear stretches of text, the word *unclear* (in angular brackets) was used and laughter was marked with the word *laughter* or any of its variants while coughing was indicated using the word *coughing* and overlapping speech was put in square brackets. Dragged words were also shown through a repetition of the dragged speech sound in the transcription, for example, a stretched *love* would be written as *loooooove*. Though this study is certainly not phonological nor prosodic in focus, a transcription of these paralinguistic and prosodic features was believed to hold a lot of potential in explaining the morphosyntactic features witnessed in ZimE.

It has been pointed out above that transcribing of oral data is time-consuming. Apart from that, transcribing different types of accent and other phonetic variations may also present challenges. Baker (2006:37) observes that “there are no set ways of responding to these issues, rather, corpus builders must reach decisions themselves, based on their own research goals. The most important aspect, however, is to be clear about the decisions made during the corpus-building stage of research and provide ample justifications for them” (Baker, 2006:37). In the context of this study, there were very few instances of speakers using strange accents. This could be a result of the fact that my research population were speakers of a variety of English familiar to me in many respects, both the population and myself being Zimbabwean black second language speakers of English. I also checked the transcriptions done by my non-Zimbabwean transcriber and corrected any misses related to accent or other linguistic differences.

For the present study, data in the spoken mode was recorded using digital audio recorders. Digital audio recorders were considered appropriate since, apart from capturing sound quality clearly, the sound files could be easily uploaded to a computer for storage and analysis. Spoken data that was used for the present study was in the form of classroom lessons, radio and television news reports and discussions, private conversations, interviews, speeches (including telephone exchanges in private and public dialogue). Data that was already in pre-recorded form was only collected and transferred to an electronic file for storage on the computer and flash discs (for back-up filing). The oral data was then transcribed with the assistance of a professional transcriber before being saved in plain text format for analysis.

3.11.3.2 The collection of written data

Before setting out to collect a written corpus, “a clear and detailed plan needs to be created” (Nelson, 2010:54). The plan should also be informed by the purpose and subsequently by issues of corpus size. Creating the plan may be a difficult hurdle to pass, but the importance of a plan lies in its role to ensure that the subsequent stages of corpus collection can be carried out and have a firm basis. Creating a plan also prevents the corpus from being simply ‘a collection of texts’ (Biber *et al.*, 1998:246). The current research used a plan in which different types of written texts to be collected were predetermined. As noted earlier, the written component of the

corpus was to be in the form of academic writing, newspaper reportage, popular writing, creative writing, student writing and letters. These different types of texts were included in order to satisfy the diversity and balance criteria that has emerged in the methodological literature. As already indicated a diverse and balanced corpus would be appropriate for purposes of adequately representing a language variety (Biber *et al.*, 1998:248; Sinclair, 2004:8).

In terms of corpus size, a written corpus can be as large as possible and to keep growing (as viewed by Sinclair (2002, in Nelson, 2010:55) but there are also other purposes to which small specialised corpora may be put. In the present study, written texts constituted 46.6 per cent of the 390 000 corpus in compliance to some extent with the conventions of the ICE model. It would have been desirable to collect a bigger component of the written corpus (and a bigger spoken corpus too) but as indicated earlier, there was need to balance corpus size with the available time and other resources.

Though there is a general belief that building a written corpus is relatively easier than building a spoken corpus, obtaining written texts to include in the corpus has been seen as the greatest challenge facing corpus developers (Nelson, 2010:60). Issues of copyright, computerising, storing arranging and cataloguing the texts are further challenges. If one has not had access or has not found pre-existing corpora appropriate for their research purposes, then one may turn to two important sources of written data; publicly available texts and privately available texts. Publicly available texts are usually in the form of newspapers, journals, magazines and internet sites (Nelson, 2010: 61). Texts obtained from an internet site however need to be cleaned of HTML formatting. Such texts also need to be checked manually to ensure, for example, that the files have not been duplicated during downloading. In terms of copyright, the corpus builder also needs to be aware of the copyright law obtaining in the country where the research is being done. For example, in the US, public data cannot be distributed as part of a corpus though it can be used to build a corpus for private research purposes.

In the current study, contrary to Nelson's (2010) observation above, the collection of the written component of the corpus indeed proved easier as most of the texts were downloaded from the internet. I sought clarification from experts on copyright laws on the use of, particularly newspaper material, for research purposes and in situations where there was uncertainty, I got

clearance from representatives of the newspaper houses concerned. These efforts were meant to ensure that no Zimbabwe copyright laws were infringed. Having been downloaded, texts that were available on the internet and other electronic sources were first saved on a computer in rich text format and then converted into plain text format. The texts were subsequently manually checked and rid of HTML formatting, images, advertisements and other irrelevant material before being annotated.

It was indicated earlier that to facilitate access, particularly to private data, personal contacts may be handy. This is because if one is personally known to the subjects, chances of gaining access to private data are enhanced. Thus, personal contacts were engaged to a large extent, not only to gain access to written texts but also to spoken data. In addition, Nelson (2010: 61) notes that there has to be “a degree of polite ruthlessness” in order to gain access to the data and that private data should also have anonymised person and company names. Efforts to maintain the anonymity of the data sources in the present study were realised through the replacement of person and company names (except for public figures) with code names in the corpus.

When it comes to data entry, Sinclair (1991:14) identifies three main methods of data entry, namely, adaptation of data in electronic format, scanning and keyboarding. A single corpus may employ multiple methods to computerise the data, though. Adaptation of data in electronic format is the easiest method though the resultant version needs to be converted into plain text which corpus analysis programmes such as *Wordsmith Tools* can process. The data that was already in electronic format, mainly newspaper reports and electronic mail was captured through such adaptation.

If data is to be scanned, then a good scanner and texts with clear typeface should be used in order to enhance accuracy. In the current study, scanning was limited to very few texts, appropriate scanners able to produce editable text not being readily available except when I visited the North-West University. The scanning method was mainly used in the capturing of the creative writing genre i.e. excerpts from novels and short stories. On the other hand, for handwritten social and business letters, as well as academic writing keyboarding was employed. Keyboarding, being the most time-consuming method is usually employed when everything else has failed and when one is converting handwritten texts or other original texts whose quality is

much degraded. All texts captured through scanning and keyboarding were manually checked, as Nelson (2010: 62) recommends that “every text has to be very carefully manually processed in order to make sure that the computer text matches the original”.

3.11.4 Nature of texts and file-naming conventions

Reppen (2010:32) observes that a key issue to be considered when creating a corpus concerns determining what constitutes a text. For this study, the corpus that was collected was organised into 195 texts each with an average of 2000 words. A single text consisted of one or more sections usually from the same source, unless untenable for some reason. So, a text in this case was not necessarily only a complete document but also an amalgamation of sections of documents into a 2000-word stretch. This was adopted mainly in genres including newspaper reportage, popular writing, student writing, letters as well as all genres of spoken discourse, because stand-alone 2000 word texts were largely hard to come by. In order to facilitate the process of corpus analysis, the texts were stored in electronic files on a computer. Backup files were stored on flash discs and external memory cards.

Reppen (2010) recommends that these files be named in a systematic way. For the purposes of this study, the texts were named according to the genre that they represented as well as according to whether they were of the oral or written mode (Adolphs & Carter, 2013:8). For example, there are file names such as “Written Creative Short Story 1” which would be the first written text in the filing system and in the form of a short story, or “Written Newspaper reportage 1” which would be the first (in the filing system) written text in the form of a newspaper report. The same conventions were used for oral texts. So there are examples such as “Spoken Public Monologue 2” which would indicate that the oral text is in the form of a public monologue and is the second in the filing system or Spoken Private Dialogue 3, which would indicate that the text is the third in the filing system in the oral mode and is in the form of a private conversation. The files were also saved in computer folders according to the genres outlined in Table 3.1.

3.11.5 Corpus analysis techniques

The critical role of computer technology in corpus linguistics is demonstrated when the corpus that has been built is now subjected to electronic processing. Evison (2010:122) remarks that this is possible because the electronic corpus is susceptible to computerised analysis using corpus access software such as *Wordsmith Tools*. Two processes that take place prior to any form of analysis are production of frequency lists and the generation of concordances (Evison, 2010:122). In addition, there are also processes of establishing collocations, normalisation as well as analysis of dispersion. Because the two processes (generation of frequency lists and concordance analysis) are used extensively in analysing the data in Chapter 4, they are examined below. (Specific descriptions in terms of the generation of concordances and related corpus analysis procedures are given in relevant sections of Chapter 4).

3.11.5.1 Frequency lists and analysis

Frequency lists capture the information regarding how many times tokens or words appear in the corpus. Frequency analysis also establishes the total number of types of each word used in the corpus. Analysis software can then display the information on a computer screen either in rank order of frequency or in alphabetical order or in reverse order (Evison, 2010:124). Frequency lists are handy for lexicographers or language syllabus designers. In terms of research, frequency analysis “can be a useful first step to get a general idea of the texts in the corpus” (Adolphs & Carter, 2013:16). Baker (2006:47) also points out other and more specific applications of frequency analysis for research purposes. For example, frequency lists “can be employed to direct the researcher to investigate various parts of a corpus, how measures of dispersion can reveal trends across texts and how, with the right corpus, frequency data can help to give the user a sociological profile of a given word or phrase enabling greater understanding of its use in particular contexts”.

Frequency analysis also includes the concept of dispersion. Dispersion has to do with the positional distribution of an item in the text i.e., does it occur more prominently at the beginning of texts, in the middle or at the end of the text? (Baker, 2006:47). Information on the dispersion

of items is important in the sense that it makes it clear that particular words are more frequent at the beginning of texts while others are more frequent at the middle or at the end of the text. Such information shows us that language is not random but follows specific structural patterns denoted by lexical patterns.

In the current study, frequency analysis plays a key role in determining the features of the variety of English used by black second language speakers in Zimbabwe as well as the variety status of the language. Reflections on the use of, for example, pronouns, verb forms, adjectives, adverbs and other lexical and syntactic structures can only be based on the frequency with which the structures are used.

3.11.5.2 Concordance analysis

The second technique that can be adopted in analysing a corpus is the establishment of concordance lines. Sinclair (1991:32) defines a concordance as “a collection of the occurrences of a word-form, each in its own textual environment” and adds that it is actually a form of index. Concordance analysis is also known as Key Word in Context (KWIC), ‘key’ meaning simply the word being examined and not necessarily a word of chief importance (Baker, 2006:71). Defining a concordance simply as “a list of all the occurrences of a particular search item in a corpus, presented within the context that they occur in; usually a few words to the left and right of the search term”, Baker (2006:71) views a concordance analysis as one of the most effective techniques to carry out a close examination of a text from a quantitative and qualitative perspective. Evison (2010:124) concurs and remarks that concordancing “is a valuable analytical technique because it allows a large number of examples of an item to be brought together in one place, in their original context.” For example, the different contextual uses of resumptive pronouns were elicited by concordancing all pronouns in English and examining each one of them to ascertain if indeed it was a case of the resumptive pronoun (See section 4.2.4). It is important here to point out that the presence of the original context here has potential to mitigate the limitation in terms of decontextualisation mentioned in 3.8 above. Baker (2006:77) goes on to observe that the essence of creating concordances is “to look for patterns of language use,

based on repetitions”, making the vertical and horizontal reading of the concordances valuable (Tognini-Bonelli, 2010:19).

A concordance analysis also elucidates a semantic preference, which is “the relation not between individual words, but between a lemma or word-form and a set of semantically related words” (Stubbs in Baker, 2006:87). Semantic preference is therefore related to collocation but unlike collocation, it focuses on a lexical set of semantic categories rather than a single word or a related set of grammatical words. Tribble (2010) explores these and many other applications and tools of concordance analysis. For the purposes of this study, it is concordance analysis that will be relied on as a window to look into, and establish, any unique ways in which morphosyntactic structures are used in ZimE.

To draw out a concordance list for resolving the extent to which a particular morphosyntactic feature was characteristic of ZimE, a word representing that particular feature would be searched for in the corpus. (As noted in Chapter 2, the candidate features for analysis were determined from the grammatical features outlined by Kortmann and Lunkenheimer (2011; 2012) as well as other information available from research on new Englishes in the African context). For example to check if there are any instances in which the pronoun *she* is used for inanimate things, the word *she* would be entered in the concordance search area and all the instances in which the pronoun is used in the corpus would appear in the concordance lines. Where the concordance lines turned out to be too many to process, they would be reduced to a maximum of 400. Then all the concordance lines would be read with the aim of taking note of all cases in which *she* was used to refer to inanimate things. Those instances that conformed to this criterion would be tagged with a particular code and the frequency of the instances established. (Please note that in Chapter 4, for each feature that is analysed, the strategy used in drawing of the concordance and all other procedures in resolving cases of the frequency of the linguistic features are described in fair detail.)

3.12 Conclusion

The present chapter has attempted to trace the history and development of what is today known as corpus linguistics or the corpus-based approach. It has explored corpus linguistics as a research method in terms of the design of corpus-based language studies, as well as practical guidelines and their theoretical motivations. This background is viewed as imperative given that the research methodology for the present study rests on the corpus linguistic approach. The chapter has also described the specific methodological activities, procedures and decisions employed for this study in terms of the accessing, capturing, filing and analysis of the texts that constitute the corpus used in the study.

Chapter 4

Data: Presentation, analysis and discussion

4.1 Introduction

This chapter presents, analyses and discusses findings regarding the notable morphosyntactic features of Zimbabwean English attested in a 390 000-word special corpus. As indicated in the literature review chapter, these features belong to grammatical categories including pronouns, noun phrases, verb phrases, negation agreement and discourse organisation and word order, as conceptualised by Kortmann and Lunkenheimer (2012). In all, eight features that were attested in the variety will be considered in detail. Table 4.1 below outlines the eight features and the morphosyntactic categories they belong to as well as the identification numbers (shown in the first column of Table 4.1) that they were given in Kortmann and Lunkenheimer's (2012) taxonomy:

Table 4.1: The eight salient features of ZimE

Feature number	Name of feature	Category
88	Extension of <i>be</i> + V-ing to stative verbs	Verb phrase: tense and aspect
174	Deletion of auxiliary <i>be</i> before progressive	Agreement
175	Deletion of auxiliary <i>be</i> before <i>gonna</i>	Agreement
194	Resumptive/shadow pronouns	Relativization
209	Addition of <i>to</i> infinitive where StE has bare infinitive	Complementation
222	Use of <i>too</i> , <i>too much</i> , <i>very much</i> for <i>very</i> qualifier	Adverbs and preposition
227	Inverted word order in indirect questions	Discourse organisation and word order
234	<i>like</i> as focussing device	Discourse organisation and word order

Section 4.2 below discusses each of these eight features under five subheadings i.e. description of the feature in standard English, possible ways in which ZimE and other varieties of English differ from standard English, corpus results, discussion and interpretation, and conclusion. The presentation of the features in the analysis follows the order in which Kortmann and Lunkenheimer (2012) list them in their taxonomy of morphosyntactic features but this does not mean that the features that are discussed first are the most prominent features of ZimE.

4.2 Salient morphosyntactic features of ZimE

4.2.1 Extension of progressive aspect to stative verbs (Feature 88)

4.2.1.1 Description of the feature in Standard English

The progressive aspect is one of the two aspects found in the English language (Minow, 2010:128) the other one being the perfective aspect. In the progressive aspect, the action is viewed as in progress or incomplete or having a limited duration (Quirk *et al.*, 1985:197; Minow, 2010:129; Huddleston & Pullum, 2002:163) while in the perfective aspect the action is viewed as having been completed. It is possible for these two aspects to combine (Quirk *et al.*, 1985:189) though.

The purpose of the progressive aspect is to indicate “a happening in progress at a given time” (Quirk *et al.*, 1985:197) and is alternatively known as the durative or continuous aspect. Huddleston and Pullum (2002:163) also distinguish between progressive aspect and progressive aspectuality; explaining that progressive aspect is a category of syntactic form while progressive aspectuality is a category of meaning. Huddleston and Pullum (2002:163) point out that progressive aspectuality entails contexts including where the situation is presented as in progress, ongoing at or throughout a specific time, where the situation is viewed imperfectly, where the time referred to is a mid-interval within the time situation and where the situation is presented as durative. Thus, the fact that the progressive denotes an ongoing action implies that it is associated with dynamic verbs. On the other hand, there are particular verbs which do typically

not take the progressive aspect in Standard English, among them stative verbs (Quirk *et al.*, 1985:198), which I describe below.

As already mentioned, the progressive aspect usually indicates that the action the verb is referring to is continuous. There are different types of progressive aspect, namely the past progressive, the present progressive, the future progressive as well as perfect forms of progressives for these three types. *The Week* of 27 June 2013 has observed that there has been expansion in the use of the progressive in the English language over the years. In Standard English, as already pointed out, a verb in the progressive aspect is usually preceded by an auxiliary *be* or any of its alternative forms such as *is*, *was*, *am*, *are* etc. Additionally, in Standard English in general, the progressive aspect of verbs is limited to action verbs and not to stative verbs and the habitual aspect (Richards & Schmidt, 2002:513) as will be shown later in this section.

The traditional conceptualisation of stative verbs is that they are verbs that denote states as opposed to dynamic verbs, which denote action (Huddleston & Pullum 2002). Richards and Schmidt (2002:513) define a stative verb as “a verb which usually refers to a state i.e. unchanging condition” and go on to give examples of stative verbs such as *believe*, *have*, *belong*, *contain*, *cost*, *differ* and *own*. In addition, stative verbs usually have to do with states of being, emotions, thoughts etc (Quirk *et al.*, 1985). Similarly, Biber *et al.* (1999:124) add that stative verbs denote “relationships and mental states of perception, cognition and emotion”. There is however a caveat that some verbs are both stative and dynamic. Examples (1) and (2) below show the word *taste* as stative and dynamic respectively:

- (1) *The juice is **tasting** great.
- (2) The connoisseurs will be **tasting** the wine just now

Similar verbs are *thinking*, *having*, *seeing* etc. Examining the traditional distinction between dynamic and stative verbs, Quirk *et al.* (1985:200-202) point out that:

Although verbs with stative meaning have been sometimes called ‘non-progressive’, we should observe that the definition of stative verbs is not so much that they are incompatible with the progressive, as that when they are combined with the progressive,

some change of interpretation other than the addition of the ‘temporary’ meaning of the progressive aspect is required.

By implication, it would be nigh impossible to establish if a given verb is stative or dynamic without digging deeper into motivating circumstances for the use of the word, a point further emphasised by Minow (2010:130) when she says that “it needs to be pointed out that verbs do not per se belong to an aspectual class” but that they are in fact used in different aspectual senses in each given context.

In addition, there are also other polysemous words such as *impress*, one of whose meanings is dynamic and can therefore take the progressive aspect while the other meaning is stative and therefore cannot take the progressive aspect. This is exemplified in examples (3) and (4) below:

(3) *She was **impressing** all the judges.

(4) They were already **impressing** patterns on the coins when the printer broke down.

Verbs used in habitual contexts, which were not attested in the ZimE corpus, also do not take the progressive aspect. In fact, making a similar point, Minow (2010:129) observes that “there are certain constraints on the use of the progressive” which require of us to appreciate a number of verb classification systems based on the lexical aspect i.e. “the inherent semantic properties of the linguistic expression used to refer to a situation” (Bardovi-Harlig cited in Minow, 2010:129) of the verb. (See Minow, 2010:128-138 for a detailed discussion of lexical aspect).

4.2.1.2 Possible ways in which ZimE and other new varieties of English differ from Standard English

As pointed out in detail in Chapter 2, there is no quantitative research on the use of the progressive aspect in stative verbs in ZimE or on any other specific feature for that matter, save a few isolated observations on the general linguistic features of English used in Zimbabwe by Magura (1984; 1985), Makoni (1993), Mlambo (2009), Schmied (1996), Ngara (1977; 1980). Acknowledging this research gap in her paper on White Zimbabwean English, Fitzmaurice (2012) indicates that she is working on the phonological and lexical features of the English of white and black Zimbabweans born after 1980. It should be emphasised, as indicated in Chapter

1 and as illustrated in Chapter 2, that a majority of the studies just mentioned relied largely on participant observation, conjecture and introspection and not on a more scientific basis such as the corpus-based method.

However, it is anticipated that speakers of ZimE will generally not have a close knowledge of the types of verbs in English to be in a position to distinguish between dynamic and stative verbs, let alone puzzle out the nitty gritty of achievement, accomplishment and activity senses of the dynamic verbs as analysed by Minow (2010:129). Stative verbs are bound to be treated as dynamic verbs and therefore it is anticipated that, despite the fact that Shona, a Zimbabwean indigenous language spoken by more than 75 per cent of the population (Ngara, 1977:41) also codes the progressive aspect, there will be a number of cases where the progressive aspect will be extended to stative verbs. This is even more likely given the background that except in Ugandan and Indian South African English, all other second language varieties of English in sub-Saharan Africa have this feature attested with either an A or B rating in submissions given by various researchers for *The Mouton world atlas of variation in English* edited by Kortmann and Lunkenheimer (2012).

Minow (2010:138-139), Van Rooy (2013; 2014) and Van Rooy and Kruger (2016) also review a number of studies done on Black South African English, among them those by Buthelezi (1995), De Klerk (2003), Gough (1996), Mesthrie (2004; 2008), Siebers (2007), Van Rooy (2006; 2011) and Wade (1997) and note that the extension of the progressive is generally fairly widespread, though the researchers differ in terms of the extent to which the progressive is used. (Some of these studies have been examined in Chapter 2). Minow's (2010) own findings show that the progressive aspect was extended to stative verbs in 11.23 cases per 1000 words, which does not vary much from De Klerk's (2003) 9.76 per 1000 though this figure is higher than Van Rooy's (2006) 6.87 per 1000 but lower than Siebers' (2007) 13.74 instances per 1000.

The mode, size and identity of the data sets for these studies could be responsible for the discrepancies in the frequencies of the progressive in Black South African English. In a paper that gives a bird's eye view of corpus linguistic work on Black South African English, Van Rooy (2013) observes that though the extension of the progressive to stative verbs is viewed as an established feature of the variety, its higher occurrence in learner English could actually indicate

that it is only a fleeting feature that is bound to dissipate with increased proficiency on the part of language users. Van Rooy and Kruger (2016) give empirical evidence to the effect that the observation referred to is indeed the case. Van Rooy (2006; 2013; 2014) also consistently argues that it is more convincing to say that the progressive is not simply extended to stative verbs but that its use in new varieties of English in general and BSAfE in particular defies the stative/dynamic limitations of verb classification and is actually an extension of the semantics of the verb.

4.2.1.3 Results from the corpus

All the instances in which the progressive aspect was extended to stative verbs were captured through running a concordance of all words ending with *-ing*. This procedure generated thousands of entries, making it necessary to randomly sample 400 entries which were checked to ensure whether or not each entry was indeed a case of the progressive aspect. The concordance lines were then manually checked to determine if the way in which the progressive was used was indeed a case of extension to stative verbs or non-stative usage. This was a necessary procedure because determining the ‘stativeness’ or ‘dynamicness’ of verbs is not always a straightforward task. As has been illustrated earlier, there are indeed verbs which may be stative and can still carry the *-ing* ending without violating Standard English conventions. Biber *et al.* (1999:472) also intimate from corpus studies that the progressive aspect is not actually limited to dynamic verbs only but actually occurs with stative verbs as well. The procedure yielded a sample of 104 progressives.

When all the concordance lines with the progressive *-ing* marker had been drawn, those lines in which there was no extension of the progressive aspect to the stative were marked with N (for non-stative progressive) and these largely included dynamic uses of verbs. The concordance lines that illustrated the extension of the progressive to stative verbs were marked with S (for extension to ‘stative’ verbs). Non-stative cases where the copula *be* had been deleted were also included. All the cases of extension of the progressive aspect to stative verbs were also checked by my supervisor and cases of disagreement were resolved jointly. The table below illustrates the

prevalence of the extension of the progressive aspect to stative verbs. Examples 5 and 6 below illustrate the extension of the progressive aspect to stative verbs:

- (5) As you might **be knowing**, lecturers downed tools mid-October up tonow (Social letters text 1).
- (6) Opposition sympathisers put it on social media showing pictures of boxes saying these were concrete evidence that they were **containing** pre-marked ballots (Spoken text 66b).

Table 4.2: Prevalence of the extension of the progressive to stative verbs

Type of progressive	Frequency	%
Non-stative use	90	86.7
Extension of Progressive to stative verbs	14	13.4
Total	104	100

It is clear from Table 4.2 that there is a considerable number of instances in which the progressive aspect is extended to stative verbs (13.4 per cent) in ZimE. This observation indicates that, ZimE deviates from Standard English conventions in its treatment of stative verbs to a considerable extent. To enable comparability with previous studies on this feature, 104 (number of instances of all progressives from a concordance of 400 lines) was normalised by calculating the total number of progressives in the entire corpus, based on the sample, and then normalised this value to instances per 1000 words, by dividing the projected total number of progressives by the total number of words in the corpus and then multiplying this by 1000. This procedure yielded 8.49 instances of the progressives per 1000 words in the corpus, of which 1.14 per 1000 words were combined with stative verbs. This result was found to be comparable with findings by De Klerk (2003), Minow (2010) and Van Rooy (2006; 2014) and to be in line with the prediction made in 4.2.1.2 above.

In terms of the distribution of the extension of the progressive aspect to stative verbs across registers, as Table 4.3 shows, there was a clear preponderance of the extended progressive in the spoken registers. This corresponds with Van Rooy's (2014) findings on the general use of the

progressive aspect in BSAfE. From the written genres, it emerged that the extended progressive occurred most with newspaper reportage.

Table 4.3: Distribution of the extension of the progressive aspect across registers

Register	Frequency	Percentage
Spoken	9	64.3
Written	5	35.7
Total	14	100

4.2.1.4 Interpretation and discussion

As indicated in the description given above, it is clear that there is a considerable number of instances in which the progressive aspect is extended to stative verbs. On the one hand, it would appear that the absence of a corresponding concept of aspect in most of the indigenous languages of Zimbabwe results in the second language users of English treating both dynamic and stative verbs the same way. However, no plausible evidence was found to suggest that this feature could be explained through direct reference to cross-linguistic influence. It would seem that, cross-linguistic influence, unlike with some features explored in this study, does not play a role in terms of how this feature is handled by second language speakers of ZimE. Just as will be observed with the deletion of the *be* auxiliary, the occurrence of the feature can be attributed in general to inadequate mastery of the second language by the Zimbabwean speakers of English. In a paper that investigates the interaction between second-language proficiency and normative processes, Van Rooy and Kruger (2016:207) observe that “candidate NE constructions that are in essence learner errors will decline in frequency in the data from increasingly proficient users partly due to teacher feedback, but conventionalised innovative constructions will remain in the data from the most proficient users, escaping censure from editors and teachers”. Such an observation can also be invoked here to explain the extension of the progressive aspect to stative verbs in ZimE though informant profile data was not a very salient parameter for analysis in this study.

More specifically, as Buregeya (2006) observes about Kenyan English, the extension of the progressive aspect to stative verbs may be a result of simplification strategies whereby second language learners do not readily recognise irregular words or syntactic patterns as such. Second language learners tend to prefer overgeneralising language rules subconsciously to obviate the obvious cognitive overload that may result from an effort to recognise all cases of irregularity in language processes and linguistic configurations. Lowenberg (1990:158) explains such nativised features in general through reference to extension of productive and morphosyntactic processes in Standard English, among them conversion of countability to non-count nouns, expansion of the lexicon through compounding, among others. As Van Rooy and Kruger (2016) argue, if such errors escape censure by teachers and editors, then they ascend into nativised features.

In addition, the attestation of the extension of the progressive to stative verbs largely agrees with what happens in other related varieties such as West and East African varieties as well as WhZimE and BISAfE which are ZimE's closest neighbours, confirming that extension of the progressive aspect to stative verbs may actually be a universal feature in African varieties of English. This also agrees with the situation in the deletion of the *be* auxiliary before verbs in the progressive aspect where ZimE tends to identify at least with East African varieties which show little or no attestation of this feature at all (Kortmann & Lunkenheimer, 2012).

However, it should be observed that instances of language use involving the progressive in general largely conform to Standard English conventions, in some cases subscribing to East African variety conventions and in other situations departing from all African English conventions i.e. West, East and Southern African English varieties. Again, this is yet another feature whose treatment provides more evidence that ZimE identifies with peer varieties in its departures from Standard English conventions.

4.2.1.5 Conclusion

Against a background where there are a number of instances showing the extension of the progressive aspect to stative verbs, it may be surmised that the feature may be given a B grade in which case it would identify with WhZimE, WhSAfE, TanzE and some West African Englishes.

Furthermore, it was also found that the instances where the progressive aspect was extended to stative verbs were clearly dominant in the spoken genres implying that indeed the feature might actually be a result of residue from second language learning processes, which might eventually dissipate with more proficiency or with editorial intervention as proposed by Van Rooy and Kruger (2016).

4.2.2 Deletion of *be* auxiliary before progressive (Feature 174)

4.2.2.1 Description of feature in Standard English

The progressive aspect or verb form belongs to the structure of the verb phrase and always works with the auxiliary *be* before the main verb ending with the *-ing* form, a form that is also used to mark the gerund-participle (Huddleston & Pullum, 2002:162-163; Quirk *et al.*, 1985:151). It was pointed out in 4.2.1, among other things, that there are informal uses in which the *be* auxiliary may be deleted. Section 4.2.1 above has detailed the characteristic functions and behaviour of the progressive, information that should be born in mind as one reads this section. I turn now to a description of the *be* auxiliary.

Quirk *et al.* (1985:120) characterise verbs in auxiliary function as ‘helping verbs’ and they divide into primary verbs (i.e. *be*, *have* and *do*) and modal verbs (*can*, *may*, *will*, *shall*, *might*, *could*, *would*, *should* and *must*). Auxiliaries make different contributions to the verb phrase e.g. contribution in terms of aspect and voice by *be* and *have*, as well as contribution in terms of modality (including such concepts as volition, probability and obligation) by modal verbs. Auxiliaries also function syntactically as operators. According to Quirk *et al.* (1985:129), the word *be* is not strictly speaking an auxiliary verb but a primary verb that functions as an auxiliary, a point also made by Biber *et al.* (1999:428). Quirk *et al.* (1985:129) go on to point out that the word *be* can be used as an aspect auxiliary or as a passive auxiliary. Expressing the same observation, Biber *et al.* (1999:428) note that “as an auxiliary verb, *be* has two distinct functions: marking progressive aspect and passive voice”. In Standard English, verbs in the progressive aspect are always preceded by a form of a *be* auxiliary though informal usage may depart from this convention.

4.2.2.2 Possible ways in which ZimE and other new Englishes may differ from Standard English

A number of new varieties of English have shown that the use of the progressive aspect of the verb without the *be*-auxiliary is very pervasive with some instances being arguably a result of faulty grasp of grammatical rules and in other instances as a marker of an informal register. Research has indicated that this feature has an A rating in Vernacular Liberian English, Nigerian English and Nigerian Pidgin and a B rating in White Zimbabwean English, Black South African English and Indian South African English.

Studies on BSAfE by Kruger and Van Rooy (2016), Minow (2010), Siebers (2013), Van Rooy (2006; 2014) and Van Rooy and Kruger (2016) make reference to the pervasiveness of this feature with Van Rooy and Kruger (2016) characterising the *be* deletion as the clearest example of a learner feature (not a conventionalised feature), among other cases involving the treatment of the progressive such as extension to stative and habitual contexts. Van Rooy and Kruger (2016:215) point out that these results indicate that the use of the progressive form changes with increased proficiency levels and that the differential rates at which the progressive construction changes provide “new insight into the interaction between proficiency and normative processes that play a role in the conventionalisation of an innovative feature or its demise as a learner error.” This background has a bearing on the explanation of reasons behind the occurrence of the feature in ZimE. Regarding the use of the progressive, it is anticipated that ZimE departs from Standard English conventions based on the observation that there is quite significant deletion of the auxiliary *be* in particular verb phrases in the progressive aspect in the neighbouring varieties, as referred to earlier. The section below reports findings from a concordance analysis focussing on the use of the progressive in the ZimE corpus.

4.2.2.3 Results from the corpus

To find the pervasiveness of the unconventional use of the progressive aspect in ZimE, the 400-entry concordance list that was used for checking the extension of the progressive to stative verbs was used, this time with the entries being coded with different symbols. Entries that used

the auxiliary *be* before progressives were coded with A (for progressive with ‘auxiliary’) and those that did not have the *be* auxiliary were coded with D (for progressive with ‘deleted’ auxiliary). All cases with the *-ing* forms that did not have the progressive meaning were marked with X and removed from the concordance. Example 7 below illustrates the use of the progressive preceded by the auxiliary *be* while Example 8 illustrates an instance where the auxiliary has been deleted:

- (7) And they are giving him or her half that two hundred dollars which is hundred dollars (Spoken text 29)
- (8) It could be another church elsewhere, or another group or people who want to call themselves a church who asking that question (Spoken text 19).

Table 4.4 below summarises the results of the concordance analysis.

Table 4.4 Prevalence of deletion of the auxiliary *be* in the progressive in ZimE

Type of progressive	Frequency	%
Progressive with auxiliary <i>be</i>	101	97.1
Progressive with deleted auxiliary	3	2.9
Total	104	100

Table 4.4 shows that there are cases (2.9 per cent) where ZimE indeed deletes the auxiliary before verbs in the progressive aspect. However, it is clear that, to a great extent, the variety largely sticks to Standard English conventions (97.1 per cent) in its treatment of the progressive aspect.

The low frequency of instances where the auxiliary *be* was used in this corpus made it unnecessary to explore the distribution of the feature in the corpus. Suffice it to say there is a marked difference in how speakers of ZimE treat progressives in different contexts. Specifically, whereas they extend the progressives to stative verbs to quite a considerable degree, the deletion of the auxiliary *be* before progressive aspect is very minimal.

4.2.2.4 Interpretation and discussion

It is clear from the rare occurrence of the deletion of the *be* auxiliary in the ZimE corpus that ZimE significantly differs from other related new varieties of English in both West and Southern Africa in its treatment of the syntactic agreement involving the *be* auxiliary. In West and Southern African varieties, the deletion of the *be* auxiliary is either very pervasive, with an A rating or neither rare nor pervasive (with a B rating) (Kortmann & Lunkenheimer, 2012). In this regard, ZimE tends to identify more with all East African varieties reported in Kortmann and Lunkenheimer (2012) i.e. Kenyan, Tanzanian and Ugandan English in which this feature is not attested at all. It is also important to note here that findings by Van Rooy and Kruger (2016) on Black South African English in which they argue that this feature is rare in Black South African English contradict Mesthrie's (2012) submissions in Kortmann and Lunkenheimer (2012).

In addition, the occurrence of the feature predominantly in the spoken corpus could be further evidence that by and large, ZimE adheres to Standard English conventions in terms of its treatment of the feature. The very minimal occurrence of the feature in the written corpus which is usually well planned and occasionally thoroughly edited is evidence that the deletion of the *be* auxiliary in ZimE is still regarded as unconventional just as is the case with how the deletion is viewed in Standard English.

Furthermore, all the instances, but one, in which the *be* auxiliary was deleted turned out to be descriptions of live events such as soccer commentaries or other descriptions of live events, as in 3 below:

- (9) Chapungu **lining** up a three-man wall as we wait to see how they are going to take this one (Spoken text 48)

In such circumstances, the *be* auxiliary would not be obligatory even in Standard English. This is further evidence that ZimE indeed adheres to Standard English conventions in terms of its treatment of the *be* auxiliary involving the progressive aspect.

A reason for the very low frequency of this feature could be that, as noted with other features, the English language in Zimbabwe is almost exclusively acquired through the school system, which system is largely informed by Standard English conventions. In a paper that assesses the

sociolinguistic status of ZimE, Marungudzi (2016a) clarifies this point. In contexts where the language is acquired through spontaneous interaction in naturalistic settings, chances are that the deletion of the auxiliary before progressives would be more pronounced. Related to this point is the fact that there were very few informal conversations in the ZimE corpus with most of the texts being formal discussions. This was a result of the fact that in ordinary conversations, English is actually a marked choice that second language speakers of Zimbabwe hardly use though code-switching abounds in such conversations.

4.2.2.5 Conclusion

This discussion has shown that there are indeed only a few rare cases where the *be* auxiliary was deleted before verbs in the progressive aspect. From this observation, it may be surmised that ZimE may have a D grade for this feature. It indeed emerged from the findings that to a very great extent, ZimE adheres to standard convention in the sense that it has an overwhelmingly huge number of instances where the verbs in the progressive aspect are indeed preceded by a form of the *be* auxiliary. Against this background, it can be concluded that this very low attestation rating could be a function of persevering interlanguage idiosyncrasies and scarcely of any cross-linguistic factors. It may also be plausible to argue that the few instances where the auxiliary is deleted may be cases of performance errors rather than any kind of feature status. This is a surprising result when compared to the eWAVE reports in Kortmann and Lunkenheimer (2012) varieties such as WhZimE and BSAfE in which the feature has a higher attestation rating as already discussed. Against this background, it may be necessary to recheck feature ratings in these varieties against concrete data. Van Rooy and Kruger (2016:210) have raised the same issue with BSAfE. Working with corpus data, Van Rooy and Kruger's (2016) findings however converge with the findings of the current study and attribute the reduction in the use of this feature to increased proficiency as mentioned earlier. It would seem that this feature is a typically West African feature where GhE and NigE have an A rating for the feature according to submissions in Kortmann and Lunkenheimer (2012).

4.2.3 Deletion of auxiliary *be* before *gonna* (Feature 175)

4.2.3.1 Description of feature in Standard English

The word *gonna* means ‘going to’, and is mainly used in informal contexts, more in speech than in writing. Biber *et al.* (1999:484) refers to this word as a semi-modal, quasi-modal or periphrastic modal, among other semi-modals like *had better*, *have to*, and *supposed to*. Biber *et al.* (1999:420) go on to indicate that this semi-modal is mainly used in British English to mark future time. Though Eastwood (1994:95) does not refer to *gonna*, he examines the standard form of *gonna* i.e. *be going to* and indicates that it is one of the many ways of indicating the future and that the choice between *will* and *be going to* depends on whether we are making a prediction about the future, expressing an intention, or talking about a plan for the future.

4.2.3.2 Possible ways in which ZimE and other new Englishes may differ from Standard English

Studies done in the context of West, East and Southern Africa show that the deletion of the auxiliary in second language varieties of English in Africa is rare or non-existent. There are actually varieties in which the feature is not attested at all, for example in all East African varieties. The feature is attested albeit with a C rating in Ghanaian and Nigerian English as well as in Southern African varieties such as White Zimbabwean English and Black South African English. Against the background that this feature has been at least attested in parts of Southern Africa, it is possible that it will be attested in ZimE. However, the feature has not been examined closely in many of these contexts in order to bring out the causes of the innovation.

4.2.3.3 Results from the corpus

To ascertain if or not there were any cases in which ZimE departed from standard convention regarding the use of *gonna*, a concordance list of all cases in which the word *gonna* was used

was drawn and checked. During the checking, all instances in which the feature was used in a conventional way were marked with C (for conventional) e.g.

(10) and I've just got one here that I'm gonna read out from Josh... (Spoken text 36).

while those that were unconventional were marked with D (for deletion (of auxiliary *be*)), e.g.

(11) And hopefully you know we gonna try and do something... (Spoken text 1)

Results of the concordance analysis are shown in Table 4.5 below:

Table 4.5: Prevalence of the deletion of auxiliary *be* before *gonna*

Occurrence of feature	Frequency	%
Conventional (C)	9	42.9
Deletion (D)	12	57.1
Total	21	100

The table shows that in all, there are 21 instances where *gonna* is used. Of these, 12 instances constituting 57.1 per cent, show unconventional use whereby the auxiliary verb just before the word is deleted while 9 instances constituting 42.9 per cent show the conventional use. This shows that the deletion of the auxiliary verb before *gonna* is a very pervasive feature in ZimE, in contrast with neighbouring varieties such as White Zimbabwean English and Black South African English.

Another important finding from the concordance analysis was that, in keeping with Standard English conventions, all the instances in which the feature occurred came from spoken texts. Even the instances in which the feature was used in a conventional way also all came from spoken texts. This confirms that the *gonna* structure is used mostly if not exclusively in oral contexts, just as is the case with standard usage (Huddleston & Pullum, 2002; Biber *et al.*, 1999).

The results from the concordance analysis also showed that most of the instances where the auxiliary was deleted occurred in sentences where the word *gonna* was preceded by a first person pronoun such as *I* or *We*.

4.2.3.4 Interpretation and discussion

It has been noted that the unconventional ways in which the word *gonna* is used in ZimE is extremely pervasive. A number of reasons could be responsible for this prevalence. To begin with *gonna* is very close in spelling and pronunciation to a Shona word *gona*, one of whose meanings is actually ‘can’ or ‘able to do something’. So when Shona second language speakers of English delete the auxiliary ‘*be*’ before ‘*gonna*’ it is actually as if they are saying “I/We/You can”, thus actually perceiving *gonna* as an auxiliary verb. For example, in (11) above, the construction could be a literal translation of:

(12) *Tine chvimbo chokuti munoziva kuti tinogona kuedza kuita chimwe chinhu...*

Against this background, it would not be necessary to have a ‘second’ auxiliary verb. However, the fact that *gonna* is only used in very few instances in the whole corpus could be a result of the fact that speakers, particularly those at the basilectal or mesolectal level may actually be under the misconception that its use is actually ‘Shonglish’, a bastard version of English in which Shona speakers tend to give Shona renditions of English sentences by literally replacing the English words with Shona words that are only phonologically related to the English words, producing a fun effect since the semantic meaning is usually totally lost. There are actually entertaining compositions of ‘Shonglish’ sentences such as:

(13) *Unodhidha ipapo waputudza here*” [You are swimming there, have you performed the necessary rituals] adulterated from “What did you put there?” or

(14) *Sofia akafa segudo*” [Sofia died in such a way as would give the impression that she was a baboon] adulterated from “So far so good”.

The second reason could actually have something to do with the nature of the word within a sentence. Read as one word *gonna* structurally looks like a word in the infinitive form and words in infinitive forms are not normally preceded by the *be* auxiliary. Standard English grammar

requires words that, instead, end with the *-ing* form or adverbs which normally take auxiliary verbs. Such a reason could actually receive additional support from the intralingual interference factor. There might also be awareness of the speakers of ZimE that the word is informal and therefore to be used only in a limited set of environments. In the final analysis, it would appear that the prevalence of the deletion of the auxiliary verb before *gonna* in ZimE is actually compounded by both interlingual and intralingual factors.

4.2.3.5 Conclusion

Against a background where ZimE evinces a very pervasive occurrence of the deletion of the auxiliary verb before the word *gonna*, in contrast to many neighbouring varieties, it may be concluded that the reason for the occurrence of this feature is most likely to be related to cross-linguistic influence whereby the word *gonna* itself incidentally happens to be a real word (of the verb class) in Shona, though spelt slightly differently i.e. *gona*. The striking difference between ZimE and other new varieties in terms of the treatment of this feature seems to offer credence to this hypothesis, though from another perspective the occurrence of the feature could be related to intralingual influence. The absence of words phonologically related to this Shona word in L1 languages of the speakers of other African varieties of English may explain the rare occurrence or non-existence of this feature in those varieties. For example, for the Shona *gona*, the Ndebele word is *kwanisa* and the Sotho equivalent is *kgona* which is very phonetically close to the Shona *gona* though Sotho has a velar affricate where Shona has a velar plosive. Though admittedly rather unusual, this finding would provide interesting possibilities for the exploration of the interplay between phonological characteristics of new Englishes users' first language and the process of grammatical feature development in new Englishes.

4.2.4 The resumptive pronoun (Feature 194)

4.2.4.1 Description of feature in Standard English

The resumptive pronoun falls under the broad category of words with a referential purpose called pronouns. Nouns also play a referential purpose. According to Biber *et al.* (1999:327), “most pronouns replace fully specified phrases and can be regarded as economy devices.” Biber *et al.* (1999) indicate that pronouns can be divided into different types including personal pronouns, possessive pronouns, reciprocal pronouns, demonstrative pronouns and indefinite pronouns, among others, but make no mention of the resumptive pronoun.

Defining pronouns as words that stand for noun phrases (contrary to traditional grammar which defines them as words that stand for nouns), Huddleston and Pullum (2002:425) discuss pronouns under the broad category of nouns and broadly classify pronouns from a functional perspective coming up with two broad categories, i.e. the anaphoric pronouns and deictic pronouns, and five narrower categories including personal, reciprocal, interrogative, relative and temporal pronouns. Huddleston and Pullum (2002:1091) make reference to the resumptive pronoun though they do not give as much detail as they give on the other types of pronouns, indicating in a footnote that resumptive pronouns are pronouns used in place of a gap in relative clauses and asserting that resumptive pronouns in English are ungrammatical. Prince (1990:482) makes a similar observation, conceding that they are not uncommon in speech though. They exemplify the use of the resumptive pronoun *they* and *them* as in (15) and (16) respectively:

(15) There are **words or terms** in this Guide that you may not be sure what they really mean.

(16) There are **words or terms** which you may not understand them

Though Huddleston and Pullum (2002:1091) note that the use of the resumptive pronoun is ungrammatical in English, they acknowledge that resumptive pronouns represent a regular feature of relative clause formation in some languages. Two African varieties that have attested this feature with an A and a B rating respectively are Black South African English and Cameroon English. However, Simo Bobda (2012) does not provide an example sentence of how the feature is used in Cameroon English. Mesthrie (2012) illustrates the use of the resumptive pronoun in Black South African English with a phrase in (17) below:

(17) ...these **things** that you call it ideophones...

Prince (1990:482) also quotes a number of examples from English to illustrate the use of the resumptive pronoun, including:

(18) There are always **guests** who I am curious about what they are going to say.

(19) The **only one** we could see her figure was Number 2.

It should be noted that all the examples of the use of the resumptive pronoun given above occur in relative clauses, in keeping with Huddleston and Pullum's (2002) conceptualisation. It however has to be pointed out that the way the resumptive pronoun configures in the context of ZimE and to some extent in BSAfE, is somewhat different from the way it occurs in the varieties mentioned above. However, the structure, according to Mesthrie's (1997:130-135) explanation of left dislocation, remains a manifestation of the resumptive pronoun in the sense that it is an extension of "the discourse function of using a pronoun to keep track of the subject (or, less commonly, another NP) when it is separated from the verb by intrusive material" (Mesthrie, 1997:132) as in 20 below.

(20) You know, all those **things** a child needs, by Maslow, Ericsson and so on and Piaget they are non-existent in the in this child eh's life (Spoken text 70).

Mesthrie (1997) goes on to give another dimension of the resumptive pronoun whereby copy pronouns in apposition with the nouns they refer to serve a resumptive function as in:

(21) **The people, they** got nothing to eat

Van der Walt and Van Rooy (2002) hold similar views, which are critiqued in Minow's (2010:67) operationalisation of the term 'resumptive pronoun'. Nevertheless, courtesy of Mesthrie's (1997) expanded structural and pragmatic conceptualisation of the resumptive pronoun, which itself is adapted from Jespersen's (1927) treatment of the phenomenon of resumption, some of the uses of the pronouns in ZimE, such as those in Examples 20 above and 22 below, among others, qualify as resumptive pronouns.

4.2.4.2 Possible ways in which ZimE and other new Englishes may differ from Standard English

Though most scholars tend to agree that the resumptive pronoun is a typically non-native phenomenon, Mesthrie (1997:133), as indicated above, notes that the feature is not completely unknown in native English varieties. Rather, the common ways in which it is manifested in second language varieties is an extension of how it is used in the native varieties.

Submissions by researchers on most African varieties of English demonstrate however that the resumptive pronoun is not a universal feature in Africa. The feature has been attested in only one East African variety i.e. Kenyan English and in all South African varieties though with varying ratings. The feature has also been attested in most Western African varieties albeit with C ratings, to a large extent. White ZimE which shares the same geographical space with ZimE has been reported not to have this feature.

Apart from a study by Wade (1996), other scholars (in the South African context) who have explored the use of the resumptive pronoun in non-native varieties of English and to various degrees and in contexts, include Botha (2012), De Klerk and Gough (2002), Mesthrie (1997; 2004; 2006), Van der Walt and Van Rooy (2002). A majority of these studies show that the resumptive pronoun is not an uncommon feature and is largely used for pragmatic purposes such as reintroducing information that has not been talked about for a while and to serve a contrastive purpose, among a myriad other. Where these purposes occur in examples examined in 4.2.4.4 below, they are explained in more detail.

4.2.4.3 Results from the corpus

A concordance analysis of the corpus revealed that there are 49 instances in which resumptive pronouns are used. The concordance also showed these instances occurred almost exclusively in the spoken corpus. As stated earlier, studies that have focused on the resumptive pronoun have proposed a number of pragmatic motivations for the use of the structure by second language speakers of English. These motivations will be explored in the context of the corpus used for this study in Section 4.2.4.4 below.

To draw all instances where the resumptive pronouns were used, a concordance list for all third-person pronouns was used i.e. *they, he, she, it, them*. It emerged from the concordance that the resumptive pronoun occurred almost exclusively with the pronoun *they* and also that all cases where the resumptive pronoun was used occurred in spoken texts.

Table 4.6 below shows the distribution of the resumptive and anaphoric pronoun *they* in a sample of 500 concordance lines.

Table 4.6: Distribution of resumptive and anaphoric pronoun *they*

Type of discourse	Resumptive (R)	Anaphoric (A)
Spoken	15	315
Written	0	170

It is evident from Table 4.6 that fifteen (15) resumptive pronouns were used out of a possible 330 contexts in the spoken corpus. It is important to note that the fifteen instances where the resumptive pronoun occurred in the corpus are all the instances that occurred in the whole sample. On the other hand, 315 instances were anaphoric instances. It is also important to note that out of a possible total of 170 chances in the written corpus, no single resumptive pronoun was used, implying that use of the resumptive pronoun in ZimE is a purely spoken phenomenon.

However, to yield a deeper insight on the distribution of the resumptive pronoun in the corpus, all the instances in which the resumptive pronouns were used were collated in order to be analysed in terms of the contexts and positions of sentences in which they occurred. As noted earlier, in all, 49 cases of the resumptive were identified in the whole corpus.

Below I discuss the occurrence of the resumptive pronoun under two topics, firstly under the distance between the resumptive pronoun and the head noun (which Mesthrie (1997)) has hypothesised, by and large, to be characterised by intervening structures such as relative clauses or other postmodifiers, and serving a variety of pragmatic functions under the phenomenon of topicalisation) as well as the semantic nature of the head nouns.

As mentioned earlier, Mesthrie (1997) has significantly expanded the structural and pragmatic scope of the resumptive pronoun. Building upon Mesthrie's (1997) observations, Table 4.7 shows the syntactic environments in which resumptive pronouns occur in the Zimbabwean corpus, particularly the word class of the words that precede the resumptive pronouns. Where the head noun was not immediately followed by the resumptive pronoun (I), four types of words were seen to precede it, namely, a hesitation marker (H), an adverb (A), noun phrases (N) and a relative clause (R).

4.2.4.3.1 Distance of resumptive pronoun to head noun

The results in Table 4.7 indicate that by far the majority of resumptive pronouns (77.5%) are preceded immediately by the head nouns they refer to, as shown in Example (22) below.

- (22) These **people** they should be named even if they refuse to comment (Spoken text 68).

This is a departure from Mesthrie's (1997) findings where there is intervening material between the resumptive pronoun and the noun phrase it refers to in a majority of his example cases. This is also a departure from reports by Prince (1990) on Standard English varieties.

Table 4.7: The immediate collocates of the resumptive pronoun in ZimE

Distance to head noun	Frequency	Percentage
I (Immediate)	38	77.5
H (Hesitation)	1	2.0
A (Adverb)	2	4.1
N (Noun phrase)	4	8.2
R (Relative clause)	4	8.2

The most likely explanation for the use of the resumptive pronoun in this context is cross-linguistic influence. This is in spite of Mesthrie's (1997:123) suggestion that in contexts like this, the resumptive pronoun is a form of focussing strategy. In Zimbabwe's indigenous languages which are largely of the Bantu family, the head noun always agrees with the following auxiliary verb or verb by means of a prefix attached to the auxiliary verb or verb. Putting it differently, Kamwangamalu and Moyo (2003:47) say in Bantu languages in general "a subject noun phrase must agree with the verb by means of an agreement prefix". For example, translated into Shona, example sentence (22) above would be:

(23) *Vanhu ava vanofanira kuzivikanwa mazita avo chero vachinge varamba kutaura chinhu*

where the *va* [they] in *vanofanira* [they must] is used to show agreement to the noun phrase *vanhu ava* [these people]. It can be argued that the second language speakers of English in Zimbabwe use the redundant pronoun *they* since it takes the place of the prefix *va* in *vanofanira*. The fact that English has no such prefixes puts the ZimE second language speaker in a strange position where he needs something to take the place of the prefix missing in the English language. Thus the Zimbabwean speakers of English assume that English should behave the way the indigenous languages like Shona, Ndebele or Tsonga do.

Resumptive pronouns that have intervening syntactic structures between them and the head nouns only occurred in contexts where there was some code-switching, such as in Example 24 below and only make up a small percentage.

(24) *Saka ah most people oh vane degree rangu they are facing problem yokuwana mabasa achona* [So ah most people oh who hold the degree that I have are facing the problem of finding jobs] (Spoken text 72).

However, it is interesting to note that the same motivation for the use of the resumptive pronoun immediately after the head noun it is referring to may be invoked for such use in this case i.e. the need to have a token word of agreement to agree with the preceding noun or pronoun. In addition, in this case, it can also be argued that the use of the resumptive pronoun serves memory-aiding purposes, particularly because the head noun will be relatively distant in temporal terms (in speech) and in spatial terms (in writing), so the listener/reader will need to be

assisted to remember what the pronoun would be referring to. The intervening language switch also seems to make the need to aid the listener's memory even more necessary.

The second highest number of resumptive pronouns are preceded in equal prevalence (4 out of 49) by relative clauses, as in Example (25) below as well as noun phrases in the form of nouns or pronouns). Examples (26) and (27) show the resumptive pronouns referring to pronouns and nouns followed by adverbials.

(25) You know, all those **things** a child needs, by Maslow, Ericsson and so on and Piaget they are non-existent in the in this child eh's life (Spoken text 70).

(26) **Some of those** they use a lot of drugs... (Spoken text 58)

(27) **Parents** sometimes of course they trust... (Spoken text 72)

There are also two instances (see Examples (28) and (29) respectively) where resumptive pronouns are preceded by adverbs and one where a resumptive pronoun is preceded by a hesitation (please note that if we took away the hesitation marker the resumptive pronoun would then be preceded immediately by the noun it refers to).

(28) Eeh **ma fees** acho [Eeh the fees themselves] they were extremely high (Spoken text 72)

(29) The **children** ah they were healthy and they didn't have any complication (Spoken text 32)

It is also important to note that there are three cases where the words preceding the resumptive pronouns are words from indigenous languages used in code switching cases, these varying from adverbs, nouns to relative clauses therefore behaving like their English counterparts.

4.2.4.3.2 Head noun semantics

Under head noun semantics, the concordance results showed that the collocates of the resumptive pronouns could be grouped into four classes. These are common nouns (C), words referring to people (P), words referring to objects (O) and proper names (N). It might be important to clarify that O class members are different from C class members in the sense that O class members comprised words that refer to inanimate things or objects and that P members are different from C in the sense that P members comprise specifically the noun *people* or its implied reference. On

the other hand, members of the C class, though a subclass of people, do not directly refer to the word *people* but narrower classes of people as in Example (30).

It emerged that the nouns that the resumptive pronouns referred to are largely common nouns that refer to subcategories of people and other animate entities. Examples (30) and (31) show resumptive pronouns referring to common nouns that sub-classify people and objects respectively.

(30) **Some of my colleagues** they go to the library (Spoken text 72)

(31) The **nicotine with the carbon monoxide** they actually work together... (Spoken text 7).

There does not seem to be something special about these nouns save that they are appropriate candidates for playing the role of subject in the sentences, in keeping with word classes that are conventionally expected to play that role. There also does not seem to be a preponderance of nouns of a particular semantic inclination though analysis using a bigger corpus could potentially generate different findings.

Table 4.8 below shows that 24 out of 49 of the resumptive pronouns referred to common nouns that refer to categories of people such as *expert*, *kids*, *management* or *politicians*.

Table 4.8: The collocates of resumptive pronouns in terms of head noun semantics

Semantics of head noun	Frequency	Percentage
C (common noun)	20	49.0
P (people)	14	28.6
O (object)	10	12.2
N (name)	5	10.2

The second highest number of resumptive pronouns are the P category most of whose members turn out to be the word *people* itself or implied reference to *people*, including the specific or implied word for *people* in indigenous languages and pronominal representatives of *people*.

Examples (32) and (33) below illustrate resumptive pronouns referring to nouns that refer to people.

(32) Because **people from the Poly** they want to use that room... (Spoken text 72)

(33) **These three** they actually work hand in glove... (Spoken text 54)

These are followed in prevalence by resumptive pronouns whose referents are nouns that refer to ordinary objects and things, as shown in Example (34) below. Example (34) is also interesting in the sense that it is part of a quotation from the book of Psalms in the Bible.

(34) **Thy rod and thy staff** they comfort me (Spoken text 84)

So it can be argued in the general analysis that resumptive pronouns occur with nouns, noun phrases or nouns with modifiers that refer to human beings in the main. This would not be surprising given that it is mainly human beings in their several subcategories as well as other animate things that are agents of the action words in sentences.

4.2.4.4. Interpretation and discussion

It has been noted that resumptive pronouns, which almost exclusively occur in spoken texts, largely occur immediately after the head nouns they refer to and that those head nouns referred to by the resumptive pronouns can be categorised into common nouns that largely refer to people, the word ‘people’ itself, proper names of people, objects and things as well as nouns with post modifiers and accusative pronouns. Various reasons for this state of affairs can be suggested.

Firstly, the fact that resumptive pronouns are predominant in spoken texts may be an indication that the use of the structure may be explained by reference to pragmatic needs of spontaneous communication. Mesthrie (1997:132) points out that the resumptive pronoun may be used to keep track of a subject or another noun phrase when it is separated from the verb by other intrusive material, as in (35) and (36) below:

(35) ... so that **people** who want to verify, they will be able to get relevant information from those people and again it gives you authority more credibility.

- (36) ...eh you know all those **things** a child needs by Maslow, Ericsson and so on and Piaget... they are non-existent in the in this child eh's life.

Thus in such contexts, the net purpose of the resumptive pronoun is to aid the listener's memory by recapitulating the subject, as it were.

Secondly, the fact that a great majority of resumptive pronouns occur immediately after the head nouns they refer to may be an indication that there are other reasons for their use apart from those given by Mesthrie (1997) i.e. that resumptive pronouns usually occur in contexts where there are intervening utterances between the head noun and the resumptive pronoun.

In addition, the semantic nature of the head noun has also shown that most of the resumptive pronouns refer to people. However most of the pragmatic functions observed to be at work in the corpus analysed by Mesthrie (1997) as well as that by Prince (1990) have not been attested in the ZimE data i.e. contrastive, listing, the 'given' function, the 'because function' etc. Thus the ZimE corpus to some degree contributes additional perspectives to the structural configuration and pragmatic functions of the resumptive pronoun. For example, we saw that there are elements of cross-linguistic influence from the structure of Bantu languages in the use of the resumptive pronoun. However, from a pragmatic point of view, no hard and fast judgments can be made as such functions tend to overlap (Mesthrie, 1997:132).

4.2.4.5 Conclusion

An interesting finding from this analysis has been that contrary to findings and conclusions made from a related corpus of South African Black English, ZimE has indeed demonstrated some significant prevalence of the resumptive pronoun but has however shown a departure in terms of the syntactic distribution of this feature. ZimE tends to have resumptive pronouns occurring immediately after the head nouns they refer to, by and large, though there are a few cases in which the referents of the resumptive pronouns are some distance from them. It also emerged that some resumptive pronouns also refer to structures in the indigenous languages of Zimbabwe e.g. *Even abadala they eh...* However, most important was the finding that virtually all cases of the use of the resumptive pronoun occurred in spoken texts showing that in such texts, the

feature may actually fall into a B if not A category. This concurs with earlier findings on the distribution of the resumptive pronoun, both in native and non-native varieties, thus giving further evidence that the resumptive pronoun plays an important conversation management role. Nevertheless, the fact that this feature occurs exclusively in spoken texts could be an indication that the feature is a result of contextual pressures associated with spoken language which is likely to dissipate in contexts with no production pressure, such as in written discourse. Otherwise, ZimE largely models itself on standard varieties of English, particularly in written discourse.

4.2.5 Addition of ‘to’ where StE has bare infinitive (Feature 209)

4.2.5.1 Description of the feature in Standard English

The infinitive falls under complementation which in general refers to the process of making a clause complete (Huddleston & Pullum, 2002). Huddleston and Pullum (2002:215) add that being more closely related to the verb (in the sense that they are dependents of the verb) complements are central elements to grammar. It is also important to point out that along with adjuncts and predicators, complements are examples of syntactic functions in clauses in the sense that they play an important part in the way words combine to form sentences (Huddleston & Pullum, 2002:44).

Infinitive complement clauses or infinitivals (according to Huddleston & Pullum, 2002) usually serve the purpose of reporting speech and mental states, reporting intentions, desires, efforts, perceptions and other general actions (Biber *et al.*, 2002: 328). According to Quirk *et al.* (1985:65), complementation is that “part of a clause or phrase which follows a word, and completes the specification of a meaning relationship which that word implies”. There are three types of the infinitive form namely the simple infinitive, the *to*-infinitive and the bare infinitive. Examples of the three types are given in (37-39) below:

(37) I better remove the furniture. (simple infinitive)

(38) You have to remove the furniture. (*to*- infinitive)

- (39) Please help them remove the furniture (bare infinitive)

In Standard English, there are a number of considerations that have to be made when attaching complements using words such as *let, make, help, dare, except/but* and *enable*.

Biber *et al.* (2002) identify five grammatical patterns in which infinitive clauses were used in their corpus, namely:

- (40) Verb + to-clause e.g. *Do not forget to close windows, I hope to finish on Sunday, or They tried to run away.*
- (41) Verb + Noun Phrase + to-clause e.g. *It will enable offenders to break free, We expect the sales to improve, or They told the learners to maintain silence.*
- (42) Verb + *for* Noun Phrase + to-clause e.g. *Finish that and wait for the assessors to arrive, Remember to ask for the dean to sign it, or They arranged for Jacob to leave early.*
- (43) Verb + bare infinitive clause e.g. *Nobody dared break the news, or, This will help reduce crime.*
- (44) Verb + Noun Phrase + bare infinitive clause e.g. *She will help mother pack the luggage, He let everybody get away with it, I saw the guests arrive, She felt her heart leap into her throat, Father will have Jan do it.*

For the purposes of this section of the analysis, it is the use of the bare infinitive clause that is of interest because it is anticipated that second language users of English in Zimbabwe would use a *to*-infinitive where Standard English has a bare infinitive. This anticipation stems from the fact that, though this feature has not been attested in either WhZimE or WhSAfE, WhZimE's closest neighbour, there is overwhelming evidence for its presence in submissions to *The Mouton world atlas of variation in English* enterprise from West, East and Southern Africa, mostly with A and B ratings.

Eastwood (1994:158) observes that there are six situations in which the bare infinitive (a type of complement clause with an infinitive verb form but without the complementizer *to*) can be used. These are illustrated in (45-50) below:

- (45) After modal verbs, e.g. The president **must** resign.
- (46) After *had better/would rather/would sooner* and *rather than*, e.g. The president **had better** resign.

- (47) After a verb plus object pattern, e.g. The tobacco **made him** sneeze.
- (48) After except/but, e.g. You have done nothing **but** play all day.
- (49) After be, e.g. All he had to do was quit the job.
- (50) In constructions such as ‘Why worry’.

Huddleston and Pullum (2002:1244) observe that bare infinitives are used after auxiliaries i.e. modal verbs (*can, dare, do, had better, may, must, need, shall, will, would rather*) and supportive *do*, the sensory perception verbs (*feel, hear, notice, observe, overhear, see*) and causatives (*have, let, and make*). Among those that are found either with or without the *to* complementizer are *ought, dare, know* and *help*. This implies that these words can be followed by a bare infinitive or a *to*-infinitive in Standard English.

After the words including *let, make, help, dare, except* and *but*, Standard English always prefers the use of the bare infinitive. Concurring with Huddleston and Pullum (2002), Eastwood (1994:158) and Biber *et al.* (2002:455) also observe that Standard English uses the bare infinitive in the context of such words as *have* when used as a causative e.g.

- (51) I will have the porter bring up your luggage

and after verbs of perception such as *see, feel* and *hear* etc. Conversely, *enable* always takes the *to*-infinitive in Standard English.

4.2.5.2 Possible ways in which ZimE and other new Englishes may differ from Standard English

Research has generally shown that second language speakers of English tend to use the *to*-infinitive in instances where Standard English uses the bare infinitive. This tendency has been witnessed in Black South African English, Indian South African English, Kenyan English and Tanzanian English, among other varieties (Kamwangamalu & Moyo 2003; Kortmann & Lunkenheimer 2012; Van Rooy 2011). It is generally believed that by and large, the treatment of complementation by Zimbabwean second language speakers of English is similar to Standard English. However, there are anticipated instances in which ZimE will use *to*-infinitives

particularly with verbs such as *help*, *dare* and *make* as well as after the conjunctive *except*. This is likely to be a result of influence generally from the Bantu languages which usually mark infinitives with a particular prefix i.e. *ku-* (Shona), *uku-* (Ndebele/XiTsonga) or *u-* (TshiVenda) usually occurring before verbs. When used in an infinitival sense, all verbs in these languages obligatorily take up the prefix. It is also expected that ZimE would behave the same way as the other English varieties spoken by Bantu language speakers such as Black South African English, Kenyan English and Tanzanian English (Kortmann & Lunkenheimer 2012) or Englishes spoken in Lesotho, Swaziland and Malawi (LSM) (Kamwangamalu & Moyo 2003). Van Rooy (2011:192) attributes any innovative linguistic features in contact situations to transfer, simplification and overgeneralisation among other similar processes.

4.2.5.3 Results from the corpus analysis

To find out how Zimbabwean second language speakers of English handled complementation using infinitives, concordance lists of a selection of the words that normally take bare infinitives i.e. *help*, *let*, *make*, *dare*, *except/but*, *have* as well as the perception verbs *see*, *hear* and *feel* were made from the corpus using Wordsmith Tools. A concordance list of the word *enable* which normally takes a *to*-infinitive in Standard English was also run to check if there were any distinctive ways in which ZimE departs from Standard English conventions on the use of the nature of the infinitive after the word *enable*. Table 4.9 below shows the prevalence of the bare infinitive occurring after *help*, *let*, *make*, *dare* and *except/but*.

Table 4.9: Prevalence of the bare infinitive in ZimE

	Frequency of bare Infinitives	%	Frequency of <i>to</i>-infinitives	%	Total
Except/But	1	14.3	6	85.7	7
Help	59	64.1	33	35.9	92
Let	14	77.3	4	22.2	18
Dare	8	80.0	2	20.0	10
Make	159	97.0	5	3.0	164
Total	241	82.5	50	17.2	291

As Table 4.9 shows, ZimE departs from Standard conventions in the use of the bare infinitives mainly in contexts where the words *except* and *but* are used. This departure has the highest percentage at 85.7 per cent. However, it should be noted that the use of this pair of words constitute the smallest total number of instances where the infinitive is used. The second highest prevalence of departure from standard convention involves contexts where *help* is used, with 35.9 per cent, followed by *let* with 22.2%. *To*-infinitives involving *dare* constitute 20.0 per cent while the most common word used with the infinitive i.e. *make* has the smallest percentage at 3 per cent. In all, it is clear that, in 17.2 per cent of cases, ZimE departs from conventional standards in the use of the bare infinitive.

To come up with a more comprehensive picture regarding the use of the bare infinitive by speakers of ZimE, a second concordance involving a few more words known to take bare infinitives was run. These words include the causative use of *have* and verbs of sensory perception including *see*, *feel* and *hear* which normally do not take a *to*-infinitive when used as perception verbs. Results of this concordance analysis are shown in Table 4.10 below:

Table 4.10: Prevalence of the bare infinitive with other words known to take bare infinitives in Standard English.

Word	Frequency with bare infinitive	Frequency with to-infinitive	Total
Have	2	0	2
Verbs of sensory perception (See/Feel/hear)	See = 6	0	6
	Feel = 0	0	0
	Hear = 0	0	0
TOTAL	6	0	6

These results indicate that the use of the three words as verbs of sensory perception in ZimE is extremely rare. For example, except for contexts involving the use of *see*, there were no cases in the corpus in which the verbs of perception with a bare infinitive were used.

Finally, a last concordance analysis involving *enable*, a word that would normally take a *to*-infinitive was run. As Table 4.11 below shows, speakers of ZimE conform to Standard English conventions in a majority of cases i.e. in 88.2 per cent of cases.

Table 4.11: Prevalence of the *to*-infinitive that conforms to Standard use in ZimE

Word	Frequency of bare infinitives	%	Frequency of <i>to</i> -infinitive	%	Total
Enable	2	11.8	15	88.2	17

4.2.5.4 Interpretation and discussion

It is clear from Table 4.11 above that second language users of English in Zimbabwe use the infinitive *to* after verbs that include *let*, *make*, *help*, *dare*, *except*, *but* and *enable* (constituting an aggregate of 17.5 per cent, which is quite significant). However, there is variation regarding the

prevalence of the grammatical patterns that are common i.e. the unconventional use of the *to*-infinitive tends to be more common when occurring after some verbs and not so common after other verbs.

As indicated earlier, instances where ZimE uses *to*-infinitives where Standard English uses the bare infinitives mostly occurred in contexts where *except/but* is used. It was observed that this unconventional structure constituted 85.7 percentage points, but there were only a total of 7 potential situations in which this structure was used, which was the smallest number involving the use of infinitives. This may lead us to conclude that, as observed elsewhere in this study, unconventional structures are most likely to occur in situations where the structures involved are not commonly used. The learners' limited exposure to the structures could be explained in terms of little or no interaction with native speakers of English who are naturally privy to exceptions to linguistic rules. The following examples illustrate the use of the *to*-infinitive after *except* and *but* respectively:

- (52) We would then be left with no other option **except** to uh boycott the payment of rates (Spoken text 28).
- (53) Of course, these are only wishes and if they could add up, then, as beggars, we have no choice **but** to ride (Popular writing 8).

It was also observed that second in prevalence in the unconventional use of the *to*-infinitive were instances involving the use of *help* (constituting more 35.9 per cent). The following four examples illustrate the use of the *to*-infinitive after *help* in different temporal contexts:

- (54) Demographic data **helps** to know the structure of the population (Student writing text 4)
- (55) And we have got two officials in the studio uh who are **helping** us to understand the laws enshrined there (Spoken text 37).
- (56) While this requirement has **helped** to determine suitability for promotion, there is still a need to interrogate this criterion to add the element of quality to the number of publications presented (Business letters 4).

It can also be argued that the significant prevalence of the considerable occurrence of the *to*-infinitive involving *help* may be due to the fact that *help* is one of the verbs which can be used optionally with a bare infinitive or a *to*-infinitive, as observed earlier (Eastwood, 1994; Mair,

1995; Huddleston and Pullum, 2002; Van Rooy, 2011). This argument would be more plausible with the further observation that the prevalence of the *to*-infinitive with *help* is followed by another verb that may also optionally take a bare infinitive or a *to*-infinitive i.e. *dare*.

It is also interesting to note that 14 instances where the *to*-infinitive is used with *help* were found in student writing, 11 in spoken texts, 9 in popular writing, 2 in creative writing and 1 in business writing. From these statistics, it can be surmised that the convention regarding the use of the *bare* infinitive by second language users of English in Zimbabwe departs from the Standard English convention principally in learner language which is still a developing language as well as spoken language where formal structures are not much of a requirement. It thus can be concluded that the affinity for standard conventions that ZimE evinces may be explained through reference to the fact that the main input of English for second language Zimbabwean speakers of English is the Standard variety which was largely diffused in the country through the education system.

The verb *let* is next in terms of prevalent use before the *to*-infinitive, amounting to 22.2 per cent of all possible situations. Example (57) below shows the use of a *to*-infinitive after *let*:

- (57) They say just because Mugabe has allowed some newspapers like the Zimbabwe Independent and the Standard to operate and also let the opposition MDC party to operate shows that he is tolerant.

As far as the use of the infinitive after the verb *dare* is concerned, it was observed that there were 20.0 per cent instances in which the *to*-infinitive was used after *dare*, with the rest being standard bare infinitives. Examples (58) and (59) are the only two instances in which the unconventional *to*-infinitive was used after the auxiliary verb *dare*:

- (58) Nobody has **dared** to...to ask or to go there

- (59) ‘Maybe it will be the only thing they get away with because it’s so heavy,’ she **dared** to hope.

This is one of the auxiliary verbs in which ZimE significantly conforms to Standard English conventions. It may be argued that the small number of instances involving the use of *dare* with infinitives (only a total of 10 instances) shows that this is a structure largely used by a few speakers who hail from the acrolectal level of the society, speakers whose command of the

language approximates that of native speakers and who are unlikely to make too many departures from the standard conventions.

Regarding the unconventional use of the *to*-infinitive with *make*, it can be observed that the reason why this was the word with which the least number of unconventional *to*-infinitives was used could be that infinitives of this word are the most commonly occurring structures in the discourse, and thus language users have had adequate occasion and opportunity to learn how to use them. Examples (60-62) illustrate the use of the *to*-infinitive with *make*.

- (60) Organisations also use internet to buy resources from other organisations **making** their business to be come easy to conduct (Student writing text 1).
- (61) “Doing gender” does not only oppress the Shangwe women, it **makes** them to feel inferior (Academic writing text 5).
- (62) The very question which **made** him to commit suicide was that so what happened to the child (Spoken text 80)

As argued elsewhere, users of English in Zimbabwe may have had enough opportunities to learn the conventional use of the infinitive involving the word, in contrast for example with the case where *except/but* and *dare* are involved.

In terms of the use of *have* the concordance results indicated that there was a very insignificant presence of the bare infinitives coming after *have* with only two instances of the use of this pattern in newspaper reportage texts. The two instances are quoted below:

- (63) If the Foreign minister needed convincing, his Cabinet colleague explained that while in reality the country was under sanctions, it should craft strategies to **have** the embargo lifted (Newspaper reportage text 13).
- (64) When we sing the "Mugabe must go, Mugabe is a murderer" song as a nation, let us have Shonas sing the verse about Gukurahundi as loud as they sing the 2008 massacres, Murambatsvina and others that affected them (Newspaper reportage text 10)

There were no instances where a *to*-infinitive was used after *have* as a causative marker.

It was noted that another set of words that normally take the bare infinitive in Standard English include verbs of perception such as *see*, *feel* and *hear* and that concordances of these three verbs showed that there were no cases in which they were followed by bare infinitives except for the

word *see* which was followed by a bare infinitive in two instances as illustrated in (65) and (66) and the word *feel* which was followed by a bare infinitive as in (67) below:

- (65) Would you want to **see** him do those now (Spoken text 16).
- (66) I **saw** little Joe and other boys scuffle over the toys (Creative writing text 6).
- (67) I **felt** them gasp and shudder and laugh as the water coursed over their strong young breasts Creative writing text 21)

Thus bare infinitives as well as *to*-infinitives involving verbs of perception in ZimE, though present are an extremely rare feature. They thus contribute very little to the overall incidence of unconventional use of the infinitive. It could be argued that this low incidence could again be testimony to the fact that the use of this feature might be a preserve of relatively sophisticated users of the language such as those from the acrolectal level as was witnessed with the use of *have* as a causative marker followed by a bare infinitive. The rest of the examples in this concordance list show that the verbs of perception are followed by a noun phrase and then a verb in the present participle. This is in itself an interesting variation from Standard English. It thus can be concluded that there are instances in ZimE where verbs in the present participle are used in situations where Standard English would have the verb in the infinitive form.

Regarding the use of the bare infinitive after the verb *enable* the concordance results show a striking difference between ZimE on one hand and Black South African English (Van Rooy, 2011) and Kenyan English (Buregeya, 2006) on the other, for example where the feature is reported to be persistent, particularly in East African English varieties and even some West African varieties such as Nigerian and Cameroon English where the feature is actually given an A rating in both cases (Simo Bobda, 2012; Taiwo, 2012). The results show that second language speakers of English in Zimbabwe virtually consistently stick to Standard English conventions when it comes to using the bare infinitive with *enable*. There were only two instances from the corpus where a bare infinitive was used after the verb *enable* and these are shown in example (68) and (69) below:

- (68) To **enable** us therefore submit our application for approval of the 2016 fees structures, we propose that we provisionally set Saturday 30 January as the date for the AGM (Business letter text 4).

- (69) ...yaa suggested some topic amendments to **enable** proceeding with what she want to focus on (Social letters text 9).

In the overall analysis, it is clear that the variety of English as used by black second language users of Zimbabwe again departs significantly from West African varieties such as Nigerian and Cameroon English which Kortmann and Lunkenheimer (2012) give an A rating. This is another feature where ZimE is closer to East African (Tanzanian and Kenyan English have B ratings for this feature).

4.2.5.5 Conclusion

It can be concluded from this discussion that the addition of the *to*-infinitive in ZimE in situations where Standard English has a bare infinitive does not necessarily cut across all possible situations where Standard English has a bare infinitive. Instead, use of the *to*-infinitive seems to be largely confined to grammatical patterns involving certain words i.e. *except/but*, *help*, and to some extent *dare*, which are typically areas where Standard English allows a choice between two variants. However, even so, preferences in ZimE tend to reflect grammatical patterns from Standard English, by and large. In a number of situations, ZimE tends to concur with Standard English, and strikingly so when it comes to the use of *to*-infinitives after the word *enable*, where there were very few instances from the corpus in which ZimE departed from Standard English conventions. Also interesting is the fact that, in terms of the choice of the infinitive type involving *enable*, ZimE has categorically differed from other related or neighbouring varieties of English such as Black South African English (Van Rooy, 2011) and Kenyan English (Buregeya, 2006) as well as West African varieties (Kortmann & Lunkenheimer, 2012). As far as the use of *to*-infinitives after words such as *make*, *dare* and *have* is concerned, the corpus contained very few instances where this pattern occurred. Against this background it can be deduced that the English variety as used by black second language users in Zimbabwe has a B grade for Feature 209, a clear contrast with WhZimE where the feature is not attested at all.

4.2.6 *Too, too much, very much* for ‘*very*’ as qualifier (Feature 222)

4.2.6.1 Description of feature in Standard English

Biber *et al.* (1999:554) discuss *very* under types of degree adverbs called amplifiers or intensifiers, adding that the role of such words is to “modify gradable adjectives and indicate degrees on a scale”. According to Carter *et al.* (2016), *very* belongs to a category of words known as qualifiers and one of the major purposes of qualifiers is to emphasise adjectives or adverbs. Carter *et al.* (2016) go on to identify some conventions regarding the use of the qualifier *very*, among them the fact that it should not be used before words in the passive voice, as in example (70) below:

(70) *Mathematics is not a **very** loved subject.

In such a case the right word to use would be *much* or *well* as in example (71) and (72) respectively.

(71) Mathematics is not a **much** loved subject.

(72) Mathematics is not a **well** loved subject.

Very may also not be used before comparatives, superlatives that use the word *most* and before ungradable adjectives. Thus we would say:

(73) This car is **much** smaller than yours, or:

(74) This car is **very much** smaller than yours,

and not:

(75) *This car is **very** smaller than yours.

Similarly, it would be unacceptable to say:

(76) *This is the **very** most memorable event.,

The acceptable form is rather:(77) This is **by far** the most memorable event

However, there is some controversy regarding the use of *very* before ‘ungradable’ adjectives such as *Christian, medical, phonological, marine* etc and ‘absolute’ adjectives such as *essential,*

equal, ideal, unique etc. Some grammarians say *very* may not be used with ungradable and absolute adjectives, but according to Huddleston and Pullum (2002:531-532), even this subset of adjectives may be qualified with words like *very* because there is a sense in which the qualities denoted by the adjectives exist in varying degrees. In other words, the adjectives can be used in both a gradable/absolute and non-gradable/non-absolute sense. Thus they argue that there would not be anything wrong with sentences like:

(78) Their behaviour is not **very** Christian.

(79) It was a **very** unique opportunity for us all.

As for the qualifier *too*, Huddleston and Pullum (2002) discuss the word under adverbial modifiers and note that the primary sense of *too* is to indicate “a higher degree than the maximum that is consistent with meeting some condition, achieving some purpose, actualising some situation...” (Huddleston & Pullum, 2002:585). The authors go on to observe that there are also conventions governing the use of *too*, a notable one of which is that a *too* structure should be always followed by a *to* or *for* phrase functioning as an indirect complement such as in (80) and (81) below respectively:

(80) He lied, knowing the sick man was **too weak to** open his mouth and set the record straight (Creative writing text 1).

(81) Zimbabwe's government is regarded way **too large for** its small economy (Newspaper reportage text 12).

However, Huddleston and Pullum (2002:585) go on to note that there are instances in which the conditions stated in the indirect complement are not always directly expressed as in (82) below, though they may always be retrieved from the context in some way:

(82) ...experts said the new Cabinet additions flew in the face of “all economic sense”, at a time that the International Monetary Fund has slashed Zimbabwe’s projected economic growth from 2,8 percent to a modest 1,5 percent this year - warning grimly that even this might prove **too** optimistic due to worsening market conditions (Newspaper reportage text 19)

Huddleston and Pullum (2002:586) further note that *too* can actually also be used in the place of *very* in informal contexts (as in (83) below) as well as in contexts where the speaker intends to mitigate a negative remark (as in (84) below :

(83) You are **too** beautiful.

(84) I was not **too** happy with the grooming.

Against this background, it was of paramount importance to scrutinise the full context of the use of the feature in order to ascertain whether it was being used in a standard or unconventional way. It was also in cases where neither the indirect complement containing the condition nor apparent implication of that condition in the context where the use of *too* would be adjudged unconventional because in such contexts *very* would be the word to use.

4.2.6.2 Possible ways in which ZimE and other new Englishes may differ from Standard English

Partly based on reports in Kortmann and Lunkenheimer's (2012) *The Mouton world atlas of variation in English*, it is likely that the English spoken by black second language users in Zimbabwe will depart from that of Standard English to a significant extent. Carter *et al.* (2016) actually note that one of the typical errors in the use of 'very' is that language users usually use 'too' to emphasise adjectives and adverbs where they should in fact use 'very'. However, it is anticipated that as has been witnessed in the other features discussed so far, these departures may not be as much pronounced as would be the case in high-contact varieties such as pidgins and creoles. It would also be anticipated that, courtesy of the geographical proximity factor, the treatment of this feature would also be influenced by, on the one hand White ZimE and on the other by other English varieties spoken in neighbouring South Africa i.e. White South African English, Indian South African English and Black South African English or at least share characteristics with the varieties. For example, Buthelezi (1995), Gough (1996) and De Klerk (2003) report this feature in Black South African English.

4.2.6.3 Results from the corpus

To determine the extent to which ZimE uses *too*, *too much* and *very much* where Standard English uses only *very*, a concordance for the word *too* was generated, complemented by a

concordance for the phrase *very much*. Once this was done, all the concordance lines were coded with letter symbols to indicate if the use of the feature was standard or unconventional. If the feature was used according to Standard English conventions, it was marked with ‘S’ (for Standard) and if it was used in an unconventional way, it was marked with ‘U’ (for unconventional). Table 4.12 summarises the results of the concordance analysis.

Table 4.12: Prevalence of *too*, *too much* or *very much* for *very* in ZimE

Qualifier	Standard (f)	%	Unconventional (f)	%
Too	51	70.8	21	29.2
Too much	11	73.3	4	26.7
Very much	80	87	12	13.0
Total	142	79.3	37	20.7

These results indicate that there is a significant extent to which ZimE departs from Standard English conventions in terms of the use of the *very* qualifier (in 20.7% of cases). There are clear instances in which ZimE uses *too* or *too much* or *very much* in the place of *very*. However, in most of these cases it is the word *too* which is used in the place of *very* (constituting 29.2 percentage points), followed by the use of *too much* for *very* (constituting 26.7 percentage points) and finally by *very much* for *very*, which makes up 13.0 per cent. These three uses are illustrated in examples (85-87) below:

- (85) I'm not **too** careful with money (Social letters text 4).
- (86) Eh eh you know what you are doing here can be **very much** scientific... (Spoken text 87).
- (87) And we have left **too much** in the sense of life (Spoken text 65).

In all, as Table 4.12 shows, the 37 instances of the unconventional use of the meaning of *very* amounts to 20.7 per cent.

All the 37 instances where the meaning of *very* was conveyed in an unconventional way were in turn checked to find out whether each feature occurred in the spoken or in the written mode of the corpus. It emerged that in an overwhelming majority of cases, the unconventional rendering of *very* occurred in the written mode. Table 4.13 presents the statistics regarding the distribution of the feature across spoken and written texts.

Table 4.13: Distribution of *too*, *too much* and *very much* as a ‘very’ qualifier across spoken and written modes

Mode	<i>Written</i>		<i>Spoken</i>	
	Unconventional	Standard	Unconventional	Standard
Nature of use				
Frequency	26	54	11	88
Percentage of unconventional use	70.3		29.7	

Table 4.13 shows that 70.3 per cent of the instances where the meaning of *very* is substituted with *too*, *too much* or *very much* occur in the written mode of the corpus while 29.7 per cent occur in the spoken corpus. However, in both written and spoken modes, the ZimE corpus evinces greater percentages of Standard use of *very* than unconventional use.

The concordance results were also further checked to determine if there were any patterns in terms of the distribution of the unconventional use of *very* in the spoken corpus and the written corpus.

Table 4.14: Distribution of *too*, *too much* and *very much* as a ‘very’ qualifier across spoken mode registers

Register	Frequency	Percentage
Public unscripted live monologue	5	45.4
Public unscripted radio interview	2	18.2
Television soccer commentary	1	9.1
Radio dj speech	1	9.1
Unscripted live workshop discussion	2	18.2

In terms of the distribution of the feature in the spoken corpus, though there were not many instances in which the feature was used, as Table 4.14 shows, a general tendency was that the feature occurred in unscripted public monologues.

As for the distribution in the written corpus, where we have the majority of instances in which *very* is used in an unconventional way, as shown in Table 4.15, the feature mainly occurred in newspaper reportage where it constitutes 34.6 per cent followed by social letters and creative writing, with 26.9 and 15.4 per cent respectively.

Table 4.15: Distribution of *too*, *too much* and *very much* as a ‘very’ qualifier across written mode registers

Register	Frequency	Percentage
Newspaper reportage	9	34.6
Social letters	7	26.9
Creative writing	4	11.5
Popular writing	3	11.5
Business letters	1	3.7
Academic writing	1	3.7
Student writing	1	3.7

4.2.6.4 Interpretation and discussion

There is clear evidence from the findings presented above that second language speakers of English in Zimbabwe deploy *very* in unconventional ways to a notable extent. It is important therefore to explore factors that could be responsible for such a scenario. One of the likely explanations for the occurrence of this feature may be related to the meanings of the words *very* and *too*. These words have overlapping meanings. According to the Carter *et al.* (2016), *very* means ‘to a great degree or extremely’ and *too* means ‘more than enough or suitable’. Furthermore, as noted earlier, Huddleston and Pullum (2002:585) opine that the primary sense of *too* is “to indicate a higher degree than the maximum that is consistent with meeting some condition”. It is clear that the two words carry concepts of ‘big proportions’ whose fine distinctions second language users of English may not be privy to. Thus the feature could be a result of the overgeneralisation and/or simplification strategies naturally seized upon by learners of second languages.

From a cross-linguistic perspective, it may also be plausible to argue that the absence of two separate equivalents for *very* and *too* in most Zimbabwean indigenous languages that are spoken as first languages in the country could be partly responsible for the prevalence of the unconventional use of the feature. For example, in Shona, the idea of emphasising verbs or adjectives is usually conveyed through the intensive extension *-sa* as in:

(88) Akanyararisa (A-ka-nyarar-isa) [*He/She is very quiet*] or [*He/She is too quiet*]

It is also important to explore the factors behind the distribution of the feature between spoken and written genres. The fact that the unconventional treatment of *very* occurred more in the written mode of the corpus shows that the feature is indeed a widely accepted one. Its use may be by no means attributed to discourse pragmatic pressures (such as limited processing time) as would apply to those that would be associated with spoken texts. Written texts including newspaper reportage, though usually produced in the context of strict timelines, and creative pieces which are among the genres in which the feature is most used, are usually thoroughly prepared and edited to ensure that they conform to conventional grammatical expectations. Now, with the assumption that such texts have undergone such scrutiny but still contain such non-

standard features, it could be concluded that these features are actually deemed acceptable in ZimE.

Focussing on the registers in which the feature occurred most, it was evident that it mainly occurred in newspaper reportage and social letters as well as creative writing. As indicated earlier on, this could be an indication that even acrolectal speakers of English who are the likely authors of newspaper reportage and fiction actually find the unconventional use of *very* as acceptable.

4.2.6.5 Conclusion

From the foregoing discussion, it can be concluded that though Black second language speakers of English in Zimbabwe largely conform to Standard English conventions in their treatment of *very*, there is clear evidence that there is a significant number of instances in which they use the feature in an unconventional way. Evidence from the ZimE corpus may suggest that this feature may be attested with a B rating. Such results may not be surprising given that this feature is among the four features that Kortmann and Lunkenheimer (2012) determine as diagnostic of African varieties of English in the sense that it is a feature that seems to be particularly African, by and large. It also emerged in the discussion that there is a possibility of cross-linguistic influence from Shona further contributed to the prevalence of the feature.

4.2.7 Inverted word order in indirect questions (Feature 227)

4.2.7.1 Description of feature in Standard English

Falling under discourse organisation and word order/syntax, questions in general refer to interrogative sentences i.e. sentences that ask questions and always end with question marks or a rising intonation in speech (Biber *et al.*, 2002:249). That questions are indeed a type of sentence is confirmed by Quirk *et al.* (1985) who, using the term ‘interrogative sentence’ discuss questions among other types of sentence that include the declarative, the imperative and the exclamative. Similarly, using the term ‘independent clauses’ for sentences, Biber *et al.*

(2002:248) classify sentences into statements, directives, questions and exclamations, indicating that “there is a general correspondence between the four basic speech-act functions and the four structural types of independent clauses”. An important characteristic of questions is that, unlike declarative sentences or statements, questions are “informationally incomplete” (Huddleston & Pullum, 2002:867). Thus, although questions can also be used for requesting, offering or suggesting, the main use of questions is to ask for information (Eastwood, 1994:26).

Questions in English divide into different types depending on the criterion used. Huddleston and Pullum (2002:865-924) use two criteria to classify and describe questions. The first criterion they use yields semantic (closed and open questions) and pragmatic (inquiries, questions intended to show surprised or interested acknowledgement of new information, indirect speech acts as well as expository questions). Another classification criterion yields seven types of questions namely: polar questions, alternative questions, variable questions, information questions, direction questions, neutral questions, biased questions, ordinary/non-echo questions and echo questions. Eastwood (1994) and Biber *et al.* (2002) identify two major categories of questions, namely *wh*-questions and yes/no questions while Quirk *et al.* (1985) identifies one more type i.e. the alternative question. There are of course as many other ways of classifying and sub-classifying questions as there are authors, which I shall not dwell on any further.

For the purposes of this study, I shall focus on *wh*-questions, one of the two main types of question discussed by Eastwood (1994:25), Biber *et al.* (2002) and Quirk *et al.* (1985:81). *Wh*-questions refer to those questions that begin with a *wh*-word “that refers to a missing element in the clause” (Biber *et al.*, 2002:250). In Standard English, there are specific conventions governing word order when formulating questions. These conventions depend on whether the question is a direct question or an indirect one. From a pragmatic perspective, direct questions which Huddleston and Pullum (2002:866) view as constituted by ‘main clause interrogatives’ refer to the asking of a question “to which one does not know the answer with the aim of obtaining the answer from the addressee”.

The word order in direct *wh*-questions is in the form: *Wh*-word + auxiliary verb + subject + main verb e.g.

(89) When are you finishing? (Social letters text 2)

An important observation to be made regarding the word order of questions is the position of the subject and that of the auxiliary verb. In direct questions, the subject normally comes after the auxiliary, which in turn follows the *wh*-interrogative form in clause-initial position. Thus *you* comes after *are* in Example (89). From general observation, it would appear that departures from the word order of English as far as direct questions are concerned are very unlikely in second-language varieties of English and therefore not being of much interest to this study. These questions will not be dealt with any further save where there arises a need to draw comparisons. I turn to indirect questions below.

Indirect questions are questions which are embedded within a higher clause, either as statements or other questions. An important distinction between a direct question and an indirect question is that a direct question is punctuated with a question mark at the end while an indirect question ends with a period. As I argue below, this is why indirect questions may not be strictly speaking regarded as questions. Indirect questions normally occur in situations where they are a complement clause of a main clause, in which case they configure as subordinate interrogatives e.g.

(90) Alright, I will explain how the falls were formed when we finish (Spoken text 61).

In terms of formal structure, indirect questions are formed by putting the general question into a subordinate clause beginning with a question word or with *if* or *whether* (Eastwood, 1994:37). It is clear from Example (90) above that an indirect *wh*-question has the word order: main clause + *wh*-word + subject + auxiliary verb + main verb, confirming that indeed subject-auxiliary inversion does not occur in indirect questions and this is because they are subordinate constructions complementing main clauses. Of course there is controversy on the concept of the structural and functional ‘subordinateness’ and ‘mainness’ of clauses in sentences (Boye & Harder, 2007; Thompson, 2002; Verhagen, 2001; and others).

In terms of pragmatic function, indirect questions make utterances less abrupt and more tentative in tone (Huddleston & Pullum, 2002:862-865). For example, it would be less abrupt to ask “Would you know what time it is?” than asking “What is the time?” all in keeping with the convention of indirect speech acts. Indirect clauses sometimes also play a factive role whereby the information given in the embedded clause is a presupposed truth. Thompson (2002:136) submits that indirect clauses serve specific epistemic, evidential and evaluative frames for the clauses they occur with. Thus from an illocutionary point of view therefore indirect questions are in fact not always real questions.

Confirming this view, Huddleston and Pullum (2002:972) point out that subordinate interrogatives, their alternative term for the complements of indirect questions, “normally express questions, but because they are embedded, there is no illocutionary force associated with them”. Thus Huddleston and Pullum (2002:862) prefer the term ‘embedded question’ to ‘indirect question’. On the other hand, though they do not make reference to ‘direct questions’, Huddleston and Pullum’s (2002) concept of ‘main clause interrogatives’ implies direct questions. According to Huddleston and Pullum (2002:972) “the main structural difference between subordinate and main clause interrogatives is that subject-auxiliary inversion does not generally apply in the subordinate construction”.

4.2.7.2 Possible ways in which ZimE and other new Englishes may differ from Standard English

A number of varieties of English have been seen to differ from Standard English in terms of their treatment of indirect questions in specific ways (See profiles of varieties of English in Kortmann & Lunkenheimer 2012). The common pattern is that users of new varieties of English maintain the word order of direct questions in indirect questions. It has been noted in the above section that when a question is embedded within another question in Standard English, the word order is supposed to change. Varieties for which this departure from standard conventions has been attested with at least a B rating include Black South African English, Indian South African English, Kenyan English, Nigerian English and White ZimE (Kortmann & Lunkenheimer, 2012). Various explanations for this non-conformity to Standard English conventions have been proffered. Examining linguistic nativisation in general, Mollin (2007:171) notes that “we may

assume that the outcome of nativisation is conditioned by processes of mother tongue language transfer, accommodation, and possibly typological constraints”. Kachru (1982) and Schneider (2003) make a similar argument.

In terms of the syntactic structure of indirect questions with reference to Bantu languages in general and to Shona, one of the major indigenous languages of Zimbabwe, in particular, in most cases there is no inversion in the indirect questions in formal contexts but informal contexts usually work with inverted word order. In fact, it is important to note that the grammatical word order behaviour of Shona is characterised by much flexibility, depending on what element of a sentence one wants to offer prominence to. Such flexibility is usually absent in the English language. The impact of these cross-linguistic differences will be explored in the section below.

It is interesting to note that since speakers of White ZimE and speakers of Black ZimE share the same geopolitical territory i.e. Zimbabwe, similarities between the two Englishes are expected. In stressing this point, Schneider (2003:243) says the settler (STL) and the indigenous (IDG) groups “who share the same piece of land increasingly share a common language experience and communication ethnography, and thus the forces of accommodation are effective in both directions and in both communities, and result in dialect convergence and increasingly large shared sets of linguistic features and conventions”. The same argument can be extended to similarities with Black South African English, Indian South African English and Kenyan English though these varieties are spoken in different geopolitical spaces.

4.2.7.3 Results from the corpus

To determine the treatment of indirect questions in the corpus of ZimE, concordance lists of all *wh*-question words (how, what, when, where, who, whom, whose, and why) were created. The concordance lists were then checked to ascertain whether all instances where these words were used were indeed instances of indirect or embedded questions. During the process of checking, all direct questions, relative clauses, cleft constructions, bare infinitive complements, exclamatives, and all instances where the *wh*-word worked as a subject were culled from the lists. These structures were taken as of no interest to the study, though they may be of interest for other studies e.g. studies that may seek to find the proportions of *wh*-word subjects to *wh*-

pronouns in the discourse of second language speakers of English in Zimbabwe. After the culling, the remaining concordance lines were checked and subdivided into those that conformed to Standard English conventions and those that deviated from the standard word order. Those that conformed to the Standard English conventions were marked with ‘S’ (for Standard) and those that deviated from the Standard use were marked with ‘I’ (for inverted word order).

Examples (91-96) show sentences in which each of the six *wh*-words were used in indirect questions in an unconventional way in the data set.

- (91) Questions we must all deal with at some stage are **when** will you retire (Popular writing Text 5).
- (92) I suggest one has to question himself first **where** would the local authorities get the money to do that whilst people are reluctant to pay their bills (Spoken text 50).
- (93) ...you asked me a question to say why eh ah **why** do people allow leaders of some churches to to to dictate (Spoken text 19).
- (94) So Nob, can you tell me **who** is Nob Styles (Spoken text 73)
- (95) His biggest mistake was that he never gave himself the space to get in touch with his true feelings, to search his heart and find **what** exactly was it that he really needed in life (Creative writing text 10).
- (96) But the question I have for them is eeh **how** many were they in the operation (Spoken text 33).

Table 4.16 shows the frequency and percentages of the each of the *wh*-question words from the concordance analysis:

Table 4.16: Concordance results on indirect questions with inverted word order

Wh-word	Standard word order (S)	%	Inverted word Order (I)	%	Total
When	1	50.0	1	50.0	2
Where	12	66.7	5	29.4	17
Who	15	71.4	6	28.6	21
Why	25	73.5	8	26.5	33
What	121	83.4	24	16.6	145
How	137	95.8	18	11.6	155
Whom	0	0	0	0	0
Whose	0	0	0	0	0
Total	311	80.8	74	19.2	385

Table 4.16 shows that inversion of word order in indirect questions features with all the *wh*-question words except *whom* and *whose* though to different degrees. In total there are 74 instances where word order in indirect questions is inverted. It is evident that inversion of word order occurs most frequently with *when*, which constitutes 50% of all instances where the word is used in indirect questions. However, it should be quickly pointed out that there were only two instances in which *when* was used in indirect questions, with one instance showing standard use and the other instance showing unconventional use. Further reflections on this scenario will be given in the interpretation and discussion section below. The second most prevalent unconventional use of the *wh*-word in indirect questions occurred with the word *where* which makes up 29.4 per cent. *Where* is followed by *who* which comprises 28.6 of all instances where the word is used in indirect questions. This is followed by *why* at 26.5 per cent. *What* and *how* have 16.6 and 11.6 percentage points respectively, and as noted earlier, there are no instances of inverted word order involving *whom* and *whose*.

4.2.7.4 Interpretation and discussion

These results show that consistently, the most frequently used *wh*-words in the corpus evince the least number of instances of word-order inversion in indirect questions. This could explain why there were 50% of indirect questions involving *when*, which is used only twice. Those *wh*-words that are least used, on the other hand, tend to have more instances of word-order inversion. This may imply that second language users of this variety have had more exposure to instances where the most frequently *wh*-words are used and this has afforded them an opportunity to master situations in which the words are used. In contrast, the least used *wh*-words tend to deprive second language users of opportunities to vicariously learn the situations in which the words should be used. In this regard, evidence from the ZimE data deviates from neighbouring Botswana English, where as mentioned later, Arua (2004) reports that inversion of word order in indirect questions is actually more widespread than the Standard English construction.

The results also show that most of the instances where there is word order inversion occurred in spoken texts. Table 4.17 illustrates that fact:

Table 4.17: Comparison of word order inversion in indirect questions in spoken and written texts

Wh-word	Frequency in spoken texts	%	Frequency in written texts	%
How	14	77.8	4	22.2
What	19	79.2	5	10.8
When	1	50.0	0	50.0
Where	4	80.0	1	20.0
Who	5	83.3	1	16.7
Whom	0	0	0	0
Whose	0	0	0	0
Why	8	100.0	0	0
Total	51	82.3	11	17.7

The preponderance of word order inversion in indirect questions in the spoken corpus may imply that, as has been observed in other features, inversion of word order in indirect questions occurs in contexts where language users seem not to have much time to self-edit their language, such as in conversation. It is clear that in the written mode where such pressures are not experienced, occurrence of unconventional language structures is reduced, the writing process by its nature affording them time to reflect on the acceptable linguistic forms expected of them. Otherwise, this is further evidence that the linguistic features of ZimE do not considerably depart from Standard English conventions.

Indeed it is clear that, by and large, ZimE conforms to the Standard English conventions in terms of word order in indirect questions. Against this background and to a great extent, ZimE may not be said to be a distinctive variety in terms of its manifestation of indirect questions. The very few instances where inverted word order in indirect questions has been witnessed may not be viewed as typical and consistent enough to warrant description as a unique feature. It may be surmised that instances of word order inversion in indirect questions may be a case of fossilised second language learner language from the less proficient users of the English language. For example, there is a villager talking to a friend in a novel who uses such constructions as:

- (97) His biggest mistake was that he never gave himself the space to get in touch with his true feelings, to search his heart and find **what exactly was it** that he really needed in life (Creative writing text 10).

There is also a case of a mesolectal speaker who makes 4 utterances such as the following in one conversation:

- (98) How do we let that situation happen] th... eh ooou you're saying **why do** people let that situation to happen (Spoken text 19).

Thus even if the indirect questions with inverted word order may be attested in ZimE, there are sobering instances where sometimes the features are concentrated in some unique speakers who may not necessarily represent the Zimbabwean user of English.

Furthermore, the fact that inversion of word order in indirect questions is a feature that cuts across traditional L1, high-contact L1, English-based creoles and indigenised L2 varieties (see Kortmann and Lunkenheimer, 2012) may mean that the explanation of the feature may be sought

in the target language. A literature search on the prevalence of this feature in the languages of Southern Africa showed that there is very little evidence to trace this feature to influences from the substrate languages of the sub-continent.

Nevertheless, Arua (2004), who focuses on Botswana English, and Fitzmaurice (2012) who focusses on White Zimbabwean English, are some of the few researchers who have explored this feature. These authors indeed acknowledge the prevalence of the feature in both Botswana and White ZimE, with the feature actually occurring more than the standard form in Botswana English in expressions such as:

(99) He asked him **what** was he looking for (Arua, 2004:265).

However, from a cross-linguistic point of view, an examination of Shona and indeed a few other other indigenous languages of Zimbabwe (as discussed below) indicate the existence of quite some flexibility in terms of the syntactic structure of indirect questions. For example:

(100) He asked me where the ball was

can be rendered in Shona as:

(101) *Akandibvunza kuti bhora raiva kupi* (literally translated to: *He/She asked me to say the ball was where*) or

(102) *Akandibvunza kwaiva nebhora* (literally translated to: *He/She asked me where was the ball*).

Again depending on what the speaker's focus is, it is perfectly in order to say:

(103) *Akandibvunza kuti bhora ndekupi kwaraiva* (literally translated to: *He/She asked me to say the ball where was it.*)

In TshiVenda, Example 100 would be rendered as:

(104) *O mbudzisa hune bola la vha hune* (literally translated to: *He/She asked where was the ball*), or

(105) *O mbudzisa uri bola lingafhi* (literally translated to: *He/She asked to say the ball was where*).

The same syntactic flexibility is also attested in Ndebele. It would appear in the final analysis that it is this inherent flexibility in terms of word order in most indigenous languages of

Zimbabwe (which would actually lead to ungrammaticality in English) that contributes to the presence of inverted word order in indirect questions. Related to this is the fact that auxiliary verbs in English are stand-alone words which can be moved around in sentences, something which is not permitted in most indigenous languages of Zimbabwe since auxiliaries usually combine to form one word with subject prefixes.

Furthermore, it is also important to note that “direct and indirect constructions are often undifferentiated in speech” (D’Arcy, 2015:46). Thus some of the questions which were deemed to be indirect from the spoken transcriptions may actually be direct questions, in which case subject-auxiliary inversion would be perfectly in order. Take for example:

(106) But one would ask why **did you** go with the police in the first place (Spoken Text 19).

This ‘indirect question’ could be rewritten as:

(107) But one would ask, “Why did you go with the police in the first place?”

On this score the preponderance of indirect questions in the spoken corpus may not be viewed as conclusive evidence that inversion of word order in indirect questions is indeed a distinctive feature of ZimE. On the contrary, it could be related to the complexities of ascertaining direct and indirect questions in the spoken mode as illustrated in the two examples above.

4.2.7.5 Conclusion

It can be concluded from this discussion that though most of the instances in which indirect questions are used by second language speakers of English conform to Standard English conventions, there is quite a considerable prevalence of indirect questions that have inverted word order, most pronouncedly with the questions involving the words ‘*when*’, ‘*why*’, ‘*who*’ and ‘*where*’. It would be predicted that word order inversion in indirect questions in ZimE may be in the C grade on average. It has been noted that the prevalence of word order inversion in this variety may be, to some degree, traceable to cross-linguistic influences though there is need for further research to establish the factors behind the use of the feature.

4.2.8 Use of *like* as a focussing device (Feature 234)

4.2.8.1 Description of feature in Standard English

In conversation analysis, a focussing device belongs to the broader category of discourse markers known as inserts. According to Biber *et al.* (1999:1086), the two broad functions of discourse markers comprise signalling a transition in the evolving progress of a conversation and signalling an interactive relationship between speaker, hearer and message. However, discourse markers are not mutually exclusive, in terms of function, apart from their roles as discourse markers; they may also function as adverbials, conjunctions or prepositions and are thus ambiguous. In broader terms, discourse markers may also function as lexical or grammatical entities. Carter *et al.* (2016) indicate that discourse markers are used to connect, organise or manage what we say during conversations or writing and express attitudes, adding that the meanings of most discourse markers are not always found in dictionaries or grammar books though their functions can be deciphered from specific discourse contexts.

Talking particularly about *like* as a focussing device, Carter *et al.* (2016) observe that *like* is usually used to draw the attention of our interlocutors to what we are going to say next and identify several other functions in which *like* may be used as a discourse marker, among them as a filler, as a strategy to soften what we have just said, and to ask for or introduce an example. Though Carter *et al.* (2016) separate these latter functions from the use of *like* as a focussing device, a close analysis of the examples submitted by researchers/informants who worked with Kortmann and Lunkenheimer (2012) in *The Mouton world atlas of variation in English* enterprise to illustrate this feature indicates that these latter functions are actually subsumed under the ‘focussing device’ function or at least the informants conceptualise them as such. Table 4.18 below shows the example sentences, the rating of *like* as a focussing device submitted by the different informants/researchers that participated in the project, as well as the varieties they reported on. The penultimate column shows my own interpretation of the pragmatic purpose of *like* as a focussing device. These purposes will be further elaborated in section 4.4 below in the context of the ZimE corpus.

Table 4.18: Example sentences of *like* as focussing device (Kortmann & Lunkenheimer, 2012)

Researcher	Variety	Example	Purpose	Rating
		SOUTHERN AFRICAN VARIETIES		
Fitzmaurice	WhZimE	You know it's <i>like</i> you go to breaktime	Filler	B
Mesthrie	BlSafE	She's <i>like</i> my cousin...	Filler	B
Mesthrie	InSafE	He's a doctor <i>like</i> .	Softener	B
Bowerman	WhSafE	I need to get <i>like</i> 90 to pass	focussing	C
		BRITISH ISLES VARIETIES		
Smith	ScE	And I had already <i>like</i> all this <i>like</i> little facts to spit out at her as soon as she said it, ken.	Filler	A
Filppula	IrE	I mean now, <i>like</i> , any home work in a big way, <i>like</i>	Filler and softener	A
Penhallurick	WeIE	its just we have so many <i>like</i> English influences around us and <i>like</i> Sian'll be by there by the side	Filler	A
Trousdale	North	How did you do that <i>like</i> ?	Ask for example	B
Wagner	SW	It was <i>like</i> moved around You had to tie your corn behind the strappers <i>like</i>	Focussing and softener	A
Trudgill	EA	That wus <i>like</i> three quid	Focussing	B
Rosen	ChIsE	It was just a big rosette <i>like</i>	Softener	B
Draskau	Manx E	Betsey was all a twitter <i>like</i>	Softener	C

It is interesting to note that the examples given presuppose a somewhat different understanding of the focussing device function on the part of the different researchers. For example, from Table 4.18 above, Draskau's example seems to show *like* being used to soften what has just been said; Trousdale's to ask for an example; Penhallurick's to function as a filler and Bowerman's and Trudgill's to 'focus attention' proper. The absence of the full context, including the oral version

or at least a ready detailed transcription of the example sentences does not help resolve this cacophonous scenario either. Notwithstanding these different interpretations of ‘focussing device,’ it is this broader conceptualisation which is adopted in this study.

Also known as ‘focuser’ (Bayley, cited in Kortmann & Lunkenheimer, 2012:161) and ‘focaliser’ (Fitzmaurice, cited in Kortmann & Lunkenheimer, 2012:489), a focussing device thus helps put specific information in focus; in other words, to give it prominence or emphasis by calling for the interlocutor’s attention beforehand. Richards and Schmidt (2002:204) note that a structure ‘in focus’ is “an element or clause that contains new information.” Biber *et al.* (1999:897) note that ‘a point of focus’ in a clause is a part of the clause which has been given prominence or emphasis, and equate ‘focus’ to nuclear stress in speech. These authors go on to note that ‘focus’ “is normally placed on the last lexical item of the last element in a clause” (Biber *et al.*, 1999:897) in conformity with the information flow principle, which states that a clause characteristically opens with given or background information and ends with new information. It follows therefore that if a phrase ends with new information, it is following the information principle that is known as ‘end focus’.

In Standard English *like* is one of a sizable number of words which have multiple meanings and uses, as already mentioned. In addition to its use as a focussing device as discussed in this section and as attested in other native varieties of English, among them English of Anglia on which most Standard English conventions are believed to be based (Kortmann & Lunkenheimer, 2012), *like* is also significantly used as a quotative device (feature 235 in Kortmann & Lunkenheimer’s (2012) taxonomy of morphosyntactic features of English).

4.2.8.2 Possible ways in which ZimE and other new Englishes may differ from Standard English

Regarding the use of *like*, ZimE and other varieties of English may differ from Standard English in the sense that there is a considerable use of *like* as a focussing device. Fitzmaurice (2012) observes about WhZimE that the use of this structure as a focussing device is widespread, though she gives it a B rating. Szrecsanyi and Kortmann (2007) also count the use of *like* as a focussing

device among what they call ‘angloversal’ features (a term that is gaining currency in the Variational Pragmatics discipline) of new varieties of English. It would therefore be expected that the use of this structure in similar ways would also be found in ZimE. From a different perspective though, given that English is acquired and used in more formal contexts in Zimbabwe and that the use of *like* as a focussing device is more of a conversational feature, there may be subdued use of the feature in ZimE. Though this feature is not widespread in East African English (probably because of little use of the English language in most contexts, courtesy of the vibrancy of Swahili), both native and non-native varieties of English in southern Africa such as Black South African English and Indian South African English on one hand and White South African English on the other, show a prevalence of this feature. Mesthrie (2012) submits an attestation B grade for Black South African English and Indian South African English while proposing a C rating for White South African English. The feature also has a B rating in White ZimE. It would also not be surprising to note that at least six varieties of the nine varieties of the British Isles attest this feature largely with an A rating (Kortmann & Lunkenheimer, 2012), making it almost an unexpected feature in new varieties of English. However, it would be interesting to find out if there are any distinctive ways in which *like* is used as a focussing device in ZimE.

4.2.8.3 Results from the corpus

To determine the extent to which *like* was used as a focussing device in the corpus, a concordance for the word was run to extract all the instances in the corpus in which the word was used. In analysing the concordance results, all instances in which *like* was used as a focussing device were tagged with F while all other instances in which it was used were marked with X. The concordance analysis yielded the results shown in Table 4.19.

Table 4.19: The prevalence of *like* as a focussing device in ZimE

<i>like</i> as focussing device	Percentage	<i>like</i> as non-focussing device	Percentage
69	10.2	607	89.8

One of the unsurprising findings from the concordance analysis, as illustrated in Table 4.20, is that *like* is used as a focussing device predominantly in the spoken corpus (constituting 87 per cent against 13 per cent for the written corpus). This is not surprising in the sense that it is in keeping with the tendency in other varieties of English both high-contact and low-contact, e.g. Mesthrie (2012), Fitzmaurice (2012), Wagner (2012), Trudgill (2012), thus confirming the fact that the feature is a principally spoken language phenomenon. Table 4.20 shows the proportion of the use of *like* as a focussing device in spoken and written texts.

Table 4.20: Distribution of *like* as a focussing device in spoken and written corpora

	Frequency	Percentage
Spoken	61	87
Written	8	13
Total	69	100

All the instances where *like* was used as a focussing device were in turn sub-classified according to the more specific pragmatic functions of the feature. To achieve this all the instances in the concordance analysis that had been coded with F were re-classified into five categories that were determined after closely studying the sentences in order to come up with the most plausible functions of the word *like* in each given context. The five categories were informed by grammatical descriptions of the word *like* in Huddleston and Pullum (2002) and Carter *et al.* (2016). Five pragmatic functions behind the use of the structure were identified. These include where *like* is used: as a vocal filler to allow for time to process thought (V), as signalling an impending clarifying example (E), as a change in the structure and direction of a conversational

turn (C), as a quotative signal (Q) and as a signal that what is following is an approximation since speaker has failed to come up with the actual phrasing (A).

Example sentences (108-112) below, drawn from the ZimE corpus demonstrate these functions, with the function given in symbol form in brackets at the end of each sentence.

- (108) They have got **like** aloe aloe vera, it's mostly found in ii in ii Tanzania and even aloe vera aa body cream... (V) (Spoken text 64)
- (109) Yes **like** actually this week, I am actually doing a private project. (E) (Spoken text 73)
- (110) Now, you have, mentioned that, ah how do you relate this issue to ah other issues related to pneumonia in terms of, **like** we are in the winter season, we are saying, we are talking of ah, the...common cold, the flus and the **like**, iii eh any relationship here? (C) (Spoken text 46).
- (111) And one example I was giving people was when I went to China, we were taken to Beijing ah museums and some of us were saying **like** what do these have to with the delegation...(Q) (Spoken text 87)
- (112) Because it becomes so adherent to the tooth surface and becomes **like** part and parcel of the tooth. (A) (Spoken text 43).

The prevalence of these functions, which are clearly not unrelated, is outlined in Table 4.21.

Table 4.21: Pragmatic discourse functions of *like*

Function	Frequency	Percentage
E	29	42.1
V	20	30.0
A	8	17.4
C	7	10.1
Q	1	1.4

4.2.8.4 Interpretation and discussion

4.2.8.4.1 Distribution of *like* as a focussing device across corpus modes

As indicated in Table 4.21 above, *like* as a focussing device is predominantly found in the spoken corpus confirming Standard English conventions, as noted before, that the use of this structure is a typically spoken mode phenomenon (Quirk *et al.*, 1985; Biber *et al.*, 1999; Huddleston & Pullum, 2002). It is important to note that all the eight instances in which *like* is not used as a focussing device come from the creative writing genre except one which comes from a newspaper report.

It is also interesting to note that most of the instances where *like* is used as a focussing device in creative writing is in the form of a conversation between characters in an informal setting, thus giving further evidence for the informal nature of this feature.

The finding that the use of *like* as a focussing device is a predominantly spoken language phenomenon in ZimE does not only resonate with findings from traditional L1 and high-contact varieties but also from sister L2 varieties (see Kortmann & Lunkenheimer, 2012; Hickey, 2012 and others).

Whereas some linguistic features discussed in this thesis were traceable to cross-linguistic influences, there can be no implication of cross-linguistic factors in this context. In fact, it could be argued that this similarity between standard varieties and non-native varieties is an indication, as we saw earlier with some features, that the English discourse practices of Zimbabwean second language users of English are actually modelled alongside those of Standard English.

However, it can be observed that this similarity between non-native and standard varieties of English would be expected in the spoken mode since most colloquial and informal language is not learnt exclusively in formal situations as is the case with the written mode of language. Actually in Zimbabwe formal examinations in English at all levels do not include oral practice, implying that even in the classroom, oral components of language are sidelined. Thus the most plausible explanation for the similarity between ZimE and Standard English in terms of the use of *like* as a focussing device would be that the Zimbabwean second language users have

absorbed Standard-like ways of using language through interacting informally with acrolectal speakers of English and to some extent with the native speakers of English themselves.

4.2.8.4.2 Distribution of *like* across pragmatic discourse functions

Table 4.21 has shown that there are a number of discourse pragmatic functions that *like* performs, among them use as a filler, use as a signal of an impending clarifying or elaborative example, as an approximation of an otherwise tricky idea to express, as a signal of a change in the structure or direction of a conversational turn or as a quotative cue.

Below is an example that illustrates the use of *like* as a filler:

(113) I will say aah you know **like**... because most of them are really good...

It is evident in this example that *like* has been used as a filler to afford the speaker time to process her thoughts. This is quite evident especially against the backdrop that the speaker has been asked a question off the cuff about who her favourite player is and she has no ready answer. Thus she buys time, not only through using *like* but also through other discourse markers such as *you know* that precede it as well as hesitation. It is also interesting that there are similar constructions whereby speakers do not just use one discourse marker item but two or several. In most instances, the focussing device *like* is actually surrounded by pauses and hesitations as the speaker gropes for words. The example below illustrates that:

(114) ...and you know that would **like**, cover, for just, maybe, just the month

The following example illustrates the use of *like* as a discourse strategy to signal an impending elaboration in the form of an example:

(115) We didn't get what you were saying, it's **like** you were reciting a poem

This kind of use is not very different from the use of examples in situations where there is no focussing function. The word is intended to signal that clarifying example. Thus discourse markers of this nature may be understood as inspired by the speaker's desire to assist the

interlocutor to get his message accurately or to process it easily, in keeping with the Gricean conversational principle of co-operation.

The next example is where *like* is used as a signal that information that is coming is not packaged in the best form of phrasing and therefore acts as a disclaimer to the interlocutor.

(116) Somehow we always you know, **like**, keep on going but we want to entertain (Spoken text 1).

Used this way, it is as if to suggest to the interlocutor that there has not been adequate time to come up with a perfect packaging of the message and therefore what he is going on to say is merely an approximation. This function may be related to the use of *like* as a filler, where the speaker is buying time in order to process thoughts.

There is also the use of *like* as a signal of a change of direction. Though this pragmatic function may be found alongside other functions such as the use of *like* as a filler, it specifically entails a need for the speaker to change course in terms of the syntactic structure and direction of the utterance. This is illustrated below:

(117) What we are doing because **like** when... when we put them to sleep we put...

It is clear in this example that the speaker has realised that it would not be possible to see his sentence through without falling into some grammatical tangle; thus he reroutes, as it were, in order to accommodate the syntactic nature of the remaining thoughts. However, it should be pointed out that such a function and indeed the rest of these purposes to which discourse markers are put are not necessarily consciously planned by the speakers. They only come to light in an *a posteriori* sort of way.

Finally there are contexts where *like* is used as part of a quotative device, exemplified by Example (118) below, though this would belong under a separate feature (Feature 235) in Kortmann and Lunkenheimer's taxonomy. As shown in section 4.3.8 below, this feature turned out to be very rare in the ZimE corpus.

(118) ...and some of us were saying **like** what do these have to do with the delegation (Spoken text 87).

4.2.8.4.3 The collocates of *like* as a focussing device

The analysis also sought to examine the use of *like* as a focussing device from a syntactic point of view, i.e. what sort of words precede and follow it? Evidence from the concordance analysis shows there is a sort of pattern regarding the kind of words that come before the focussing device *like*, with auxiliary verbs *it's* and other verbs topping the list. With regards to the words that come immediately after the device, there was no clear pattern in terms of the kind of company that the focussing device keeps it emerged that there were all classes of words though pronouns and determiners were in the majority.

4.2.8.5 Conclusion

Findings from this section indicate that there is hardly any unique way in which ZimE deploys *like* as a focussing device in communication. Evidence from the concordance analysis shows that the feature is used largely in the same way in which other varieties, regardless of whether they are high-contact or low-contact varieties treat it. Since there are similarities in the treatment of this feature even between WhZimE and WhSafE on one hand and ZimE and B1SafE on the other, the only plausible explanation for the treatment of the feature would be influence from the white settler population. It is somewhat puzzling that this feature is not attested in any East African variety, varieties that are believed to be premised on British Standard English (Kortmann & Lunkenheimer 2012). This could point to the fact that spoken ZimE is more widespread than East African Englishes. In the final analysis, it can be argued that the treatment of *like* as a quotative device, free from cross-linguistic influence as it seems, is further evidence of the close relationship between ZimE conventions and Standard English conventions.

4.3 Low-frequency/Unattested features

This section focuses on those features that yielded very low frequencies or features that did not occur at all in the ZimE corpus. Among the features are the four identified by Kortmann and Lunkenheimer (2012) as particularly African features. The rest of the features were determined through a comparison of ZimE with related second language varieties, especially in Southern Africa. Among the low-frequency or unattested features are also features that Kortmann and

Lunkenheimer (2012) report to be absent in Africa for which it was necessary to confirm using the ZimE corpus. The low-frequency or unattested features include the following:

4.3.1 *She/her* used for inanimate referents (Feature 1)

Example 119 below illustrates the use of *she/her* to refer to inanimate things

- (119) Disclosures of the alleged secret meeting between the two leaders, is coming on the backcloth of reports that **China** has "demanded clarity" on the sensitive matter, and has stepped up financial and political support for their desired candidate, Mnangagwa, who received part of his military training in China during the 60s, as the Asian powerhouse angles to secure **her** wide economic interests in Harare through a trusted politician with close links to Beijing.(Newspaper reportage text 10).

This feature is reported by Kortmann and Lunkenheimer to be attested in all other English varieties except African varieties. In the context of ZimE, to check for the occurrence of the feature, a concordance for *she* and *her* was generated and it emerged that there was only one instance in which the feature was used from a 400-entry concordance sample. According to Kortmann and Lunkenheimer, such a low frequency would still fall under the absence of feature category and in this respect, ZimE is like its peer varieties in its avoidance of *she* for inanimate things.

4.3.2 *Me* instead of *I* in coordinate subjects (Feature 7)

To identify the instances where *me* instead of *I* was used in co-ordinate subjects, I concordanced for *me* and *I*, taking note of those particular instances in which either of the pronouns was in a co-ordinate relationship with another noun phrase. Though a concordance list run for this feature produced six instances in which *me* or *I* were in a coordinate relationship with other noun phrases, only one (shown below) of these showed that the coordination was non-standard.

- (120) "There are guys like **me and Kieran** who have played at that level before and then there are the younger guys who I'm sure will enjoy the experience of playing team golf rather than individual golf," he said.

Clearly, this is yet another feature where the adherence by ZimE to Standard English conventions comes out quite strongly.

4.3.3 *Meself/myself* instead of *I* in coordinate subjects (Feature 8)

To identify all the instances in which *meself/myself* was used in coordinate subjects, I concordanced for the words *meself* and *myself* respectively. The results showed that the corpus contained only 54 entries in which *myself/meself* was used and only two of these (*myself* in both cases) were used in co-ordinate subjects i.e.

- (121) **Your SDC and myself** will say more about this grand acquisition next year January, God willing, when you come for the AGM. (Business letter text 4)

The other case, i.e.

- (122) More importantly, I believe in finding common ground for the sake of **myself and my fellow young people** (Popular writing text 10).

is interesting in that *myself* even comes first, making it doubly unconventional, i.e. in terms of position and in terms of form.

Though this frequency is clearly very low, it is striking that in the two cases in which *myself* occurs in a coordinate subject, it is used in an unconventional way. It is also striking that there is no occurrence of the *meself* variant of *myself*. It may be suggested here that this could be another feature that could be explored further with a bigger corpus. It could also be said that, just as with the feature above, the reason why the feature occurs has to do with the very low frequency of the constructions in general use. Such a scenario would inevitably prejudice language users of exposure to the structure (as Kruger & Van Rooy (2016) argue in another context) resulting in the language users having very limited opportunities to learn to use it in a conventional way.

4.3.4 Different count or mass noun distinctions resulting in use of plural for StE singular (Feature 55)

A concordance list was run to check the occurrence of all words that are usually pluralised erroneously for example *furnitures* for *furniture*, *equipments* for *equipment*, *advices* for *advice*, *luggages* for *luggage*, *informations* for *information* and so on. These were sampled from English grammar books, especially Carter *et al.* (2016) and Huddleston and Pullum, (2002). It turned out that there were very few instances where such a feature is used in ZimE despite the fact that this is one of the commonest features in other African varieties of English according to Kortmann and Lunkenheimer (2012:811). Only one unconventional case of the non-count noun '*equipments*' was found, namely,

(123) What we had was two sets of equipments in the theatre (Spoken text 33)

There were also '*waters*' and '*foods*' but these seem to have been largely used conventionally. For example, out of twenty cases of '*foods*' only one seemed unconventional, namely,

(124) ... we found that the current eating patterns suggest that eh... the foods include cooking oil and all of us we need cooking oil to prepare our food especially vegetables. (Spoken text 42).

'*Fruits*' was also used four times but there was only one that was used unconventionally, namely,

(125) Currently, Zimbabwe is focusing on importing fertilisers, building materials, spare parts and equipment from Russia, while it exports mainly tobacco, **fruits** and souvenirs to Russia.

4.3.5 Double determiners e.g. demonstrative/article + possessive, with possessive pronoun preposed or postposed (Feature 59)

This feature was searched for through concordances for all permutations of determiners and possessives. This was possible because these constructions belong to classes that have a finite number of members. The only permutation that yielded some results was of 'this' and 'our' with 'our' postposed to form '*this our*' as in;

- (126) We believe that nothing at this juncture should come up and stop us from exercising **this our** constitutional right to demand our rights.

However, in both of these cases, the construction was used in a standard way. According to Kortmann and Lunkenheimer (2012:811), along feature 55, discussed above, this is one of the four features attested in Africa, which is considerably less common in the world, and can therefore be said to be particularly African. Again here, ZimE departs from its peer varieties to adhere to Standard English conventions.

Kortmann and Lunkenheimer (2012:811) also note that there are about 73 features that are not found in African varieties and such features also negatively define the varieties. Thus, this feature can be used to negatively define ZimE.

4.3.6 Extension of progressive to habitual contexts (Feature 89)

It should be noted that there are various contexts in which the extension of the progressive in habitual contexts is acceptable (Kruger & Van Rooy, 2016:11; Quirk *et al.*, 1985:199; Sharma, 2009:180; Van Rooy & Kruger, 2016:213). According to Quirk *et al.* (1985:199), the extension of the progressive aspect to habitual contexts is acceptable in contexts of repetition over a limited period, where a sequence of events has duration or incompleteness, when used with *always*, *continually* or *forever* and to show informality, usually in letter writing (such as *I am writing to invite you...*), are said to be standard.

For the purposes of this study, extension of the progressive aspect to habitual contexts entails the use of the habitual in non-delimited contexts (Sharma, 2009:180), such as

- (127) I am eating meat (Sharma, 2009:180) (meaning ‘I eat meat as part of my usual diet’).

What clearly makes this construction non-standard is the fact that there is no time-bound phrase that Quirk *et al.* (1985:199) refer to, to licence it as a standard construction. Thus such a construction becomes a good example of non-standard extension of the progressive aspect to habitual contexts.

Though two other features (extension of the progressive aspect to stative verbs and deletion of the auxiliary *be* before verbs in progressive aspect) related to the treatment of the progressive aspect were attested in the ZimE corpus, there were absolutely no cases of extension of the progressive form to habitual contexts. Entries from the concordance file that had been used to explore the extension of the progressive aspect to stative verbs were re-coded to capture habitual and non-habitual uses of the progressive form. It emerged from the analysis that all the instances of the extension of the progressive aspect to habitual contexts were within standard conventions or undecided i.e. they depicted the internal temporal structure of a recurrent event such as in;

(128) The professor is typing his own letters while his secretary is ill (Quirk *et al.*, 1985:192).

This finding echoes that of Fitzmaurice (2012) on WhZimE. Researchers on related South African varieties of English also report a low attestation of this feature. However, on the whole, it is also important to note that ZimE clearly departs from all its peer varieties (in the West, East and South of Africa) sticking closer to WhZimE and WhSAfE, together with Tanzanian English.

4.3.7 *Come*-based future ingressive markers (Feature 116)

Exemplified in Kortmann and Lunkenheimer (2012) with the construction ‘I am coming to cook your meal’ to mean ‘*I am about to cook your meal*’, this is the only one of the four typically African features that was not attested in the Zimbabwean corpus at all. To check the occurrence of the feature, a concordance of the word *coming* was run on the corpus and all instances of the word did not have the future meaning implied in Kortmann and Lunkenheimer’s (2012) example. Therefore, this is yet another feature that can be used to negatively define ZimE.

4.3.8 Use of *like* as a quotative device (Feature 235)

Though this feature was anticipated to be pervasive in the ZimE corpus (partly because of its attestation in related African varieties and also through the researcher’s own intuition) a concordance for the word *like* which was generated to identify the occurrence of *like* as a

focussing device revealed that there was only one instance in which *like* was used as a quotative device i.e.

- (129) ...we were taken to Beijing ah museums and some of us were saying **like** what do these have to with the delegation for...

This example is different from the rest of the examples given in submissions to Kortmann and Lunkenheimer's (2012) for *The Mouton World atlas of variation in English* in that the quotative *like* is actually preceded by *saying* which makes it rather redundant since the two words actually have the same pragmatic function.

As argued in 4.2.8 above, it appears that the use of *like* as a quotative device is indeed a subclass of the broader focussing device category. This feature has also been reported to be a characteristic of informal conversation particularly among the youth (Kortmann & Lunkenheimer, 2012) and against that background, recourse to a more specialised corpus, such as one made up exclusively of informal conversation, is likely to yield a much wider attestation of this feature.

4.3.9 Conclusion

It would be concluded that these features could be used to negatively define ZimE. However, the fact that the absence or low frequency of the features was based on on a somewhat small corpus would make it necessary to verify such a conclusion with a bigger corpus. Most research on varieties of English has exclusively concentrated on features with higher frequencies in the data set but a focus on features that are absent and explanations behind those will certainly complement our knowledge of these varieties so far. It would also be interesting, as Kortmann and Lunkenheimer (2012) submit, to continue looking for any other variety features which are not listed in the taxonomy proposed in *The Mouton atlas of variation in English*.

4.4 Variety status of Zimbabwean English

Corpus evidence from ZimE as discussed in 4.2 above shows that, though somewhat subdued, there is clear evidence of nativisation. It should be emphasised that, as consistently indicated in this thesis, conferment of variety status to a new English is not only warranted by linguistic character, notwithstanding the central role of this factor. Against this background, ZimE can be said to be a new English albeit one which has stalled at the nativisation stage in its development due to the absence of codification and acceptance of the variety, as will be argued below. In other words, it cannot be said to have achieved autonomous ontological status. This is not helped by the fact that a considerable number of characteristics attested in the ZimE corpus collected for this study are considerably in keeping with the conventions of Standard English. Furthermore, almost all the characteristics attested in pidgins and creoles, which arguably have the remotest relationship with Standard English, are absent from ZimE. This finding, through corpus evidence, is in agreement with that of Makoni (1993:97) where he observes that ZimE is only an interlanguage and not a new English because “there are powerful sociolinguistic and practical factors in Zimbabwe which militate against the standardisation of an interlanguage in the presence of a more prestigious, full-fledged native-speaker variety”. As pointed out in Chapter 2, in terms of Schneider’s (2003; 2007) Dynamic Model, Zimbabwe may actually be at stage 3 of the model since there is definitely evidence of structural nativisation with regards to some features. However, as we saw in Chapter 2 regarding the determination of the status of new Englishes, settling the variety status question is not always a straightforward task as there are various factors, the sway of each of which has not been so far determined in order to settle the question with a modicum of finality. There are questions as to whether the structural features concerned are errors or conventions. According to Bambgose (1998) as discussed in section 2.5 two factors i.e. codification and acceptance, a part of institutionalisation according to Mollin (2007), play a critical role in the determination of whether or not a linguistic form is an error or an innovation. When we look at the linguistic features discussed in 4.2 above in terms of codification, it can be seen that there is no evidence of codification of ZimE as I argue elsewhere (see Marungudzi 2016a). Thus, though there is clear evidence of nativisation in the use of some features, the features have not yet been codified as evident in the absence of dictionaries, grammar books and usage guides on ZimE, and are thus still viewed as errors. This is despite the existence of an English Language reader titled *English for Zimbabwe* which has been used in

Zimbabwean secondary schools since the 1980s. The conspicuous absence of evidence of the critical process of codification of ZimE clearly demonstrates that we do not have in place the label “Zimbabwean English” in the sense of “X English” proposed in Schneider’s (2003) model.

Furthermore, that there are only a few morphosyntactic features that would point to the existence of ZimE as an independent variety of English is somewhat atypical given that the extensive use of the English language in all the key sectors of Zimbabwean life would in fact constitute fertile ground for the development of a unique variety. Indeed, from a sociolinguistic point of view it would appear that ZimE is almost an autonomous variety of English, and as mentioned earlier, at the phonological level, it is an autonomous variety (Kadenge, 2009; 2010). Marungudzi (2016a) makes reference to the widespread use of the English language in sectors such as the media, education, science and technology, law, business and government administration to demonstrate why ZimE would deserve the description of ‘new English’ from a sociolinguistic point of view. More curious would be the fact that the late 1990s have seen flaring unprecedented anti-British sentiments (Ndhlovu, 2011) and a vigorous if not hypocritical drive towards indigenisation of almost all aspects of Zimbabwean life, among them the local content mantra on ZBC TV, a basis on which the public broadcaster’s current slogan is “*The pioneers of local content*”.

It can therefore be predicted that as globalisation advances and there is more interaction through social media platforms that are connecting people and enabling them to share ideas, and as the attitudes of Zimbabweans towards English in general and towards Standard English in particular continue to be positive, it is unlikely that a conventional and autonomous Zimbabwean variety of English will emerge. Recent research has also pointed to not so many families using the English language as a home language. Instead, as Schneider (2014:28) explains in the transnational attraction model, ZimE is likely to be influenced by global language change trends. This will not be helped by the fact that a considerable size of the Zimbabwean population now lives and work outside the country, with South Africa alone estimated to be hosting between 1 million and 3 million Zimbabweans (Chiumia, 2013).

The situation will also be compounded by the unmitigated exodus of white native-speaking Zimbabweans since the beginning of the land reform programme of the late 1990s (in the sense that the STL strand, an invaluable factor in the development of a new variety would have been

removed), though from a different perspective this could actually promote the birth and development of a unique Zimbabwean variety of English (in the sense that there is no more checking or moderating effect from the STL strand, therefore the local variety can actually develop with little or no interference, as it were).

It has been pointed out that the main vehicle for the diffusion of the English language in Zimbabwe is education. It has been consistently noted that in Zimbabwe the education and examination system is based on external standards (Makoni, 1993; Mareva *et al.*, 2016; Marungudzi 2016a; Mlambo, 2009). The impact of this on the development of ZimE is two-fold. Firstly, this practice explains, as already seen earlier, why there is little departure from Standard English conventions. Secondly, there is very little use of English outside formal settings, a factor which would play a significant role in the development of the variety. It has already been explained that the use of indigenous languages of wider communication for inter-ethnic communication in a way helps reduce the functional space of English, thus compromising its potential to develop as a new English. Of course we have code-switching thriving in many contexts (See for example Dube & Cleghorn, 1999; Mareva & Mapako, 2012; Marungudzi, 2014; Mugari, 2014 and Viriri & Viriri, 2013).

In other words, Zimbabwe clearly carves a diglossic context in which there is widespread use of English in the formal domains and widespread use of indigenous languages, even with code switching and code mixing, in informal situations.

In the final analysis it can be argued that in terms of Schneider's dynamic model of the birth and development of new Englishes, ZimE is partly at the exonormative stage and partly at the nativisation stage. The exonormative aspects can be attributed to the fact that, despite aspects of nativisation, there is corpus evidence pointing to the reliance of variety still on external norms i.e. Standard English. Because new English in the sense of "X English" seems to apply only when the variety has reached the endonormative stage (stage 4), strictly speaking, because of the absence of institutionalisation of the variety, ZimE would therefore not be viewed as a new English, with little prospects of progressing beyond the nativisation stage.

4.5 Conclusion

In the final analysis, it is clear that there is a sizable number of features that characterise ZimE though there do not seem to be features that are exclusively unique to this variety except for feature 175 (deletion of be auxiliary before *gonna*). It also emerged from the discussion that ZimE largely identifies with its neighbouring Southern African varieties of English though in respect of certain features it may, in addition, also identify with West and East African varieties. On the other hand, there does not seem to be much overlap between ZimE and non-native speaker varieties as well as pidgins and creoles, with ZimE showing a clear tendency towards Standard English. All these points certainly detract from the potential configuration of ZimE as an autonomous variety of English. Similarities between ZimE and other second language varieties of English result in an interesting conclusion about the causes of the notable features of the variety, namely that there may not be a single factor behind the occurrence of any certain linguistic feature. Actually as the discussion unfolded it became more and more apparent that one cannot put a finger on only one factor such as cross-linguistic influence, second language learning strategies, pragmatic expediency/constraints to explain the occurrence of the feature. It turns out that motivations behind the use of particular features may largely be a matter of conjecture and it seems the inclusion of closer biographical information of language users as well as an examination of the contextual factors under which particular discourses are produced, which can be enhanced by the use of multimodal data, has potential to open windows into explanations of the features of new Englishes.

Chapter 5

Conclusion

5.1 Introduction

As a lesser-known variety of English, the English spoken in Zimbabwe has not received much attention from researchers. The variety, along with some of its other southern African peer varieties, has been mistaken to be part of the South African varieties. Though scholars such as Ngara (1977; 1980) have looked at the phenomenon of English in Zimbabwe, attention to ZimE in the context of the study of new Englishes began with Magura (1984; 1985) and Makoni (1993), before being taken up by Mlambo (2009) as well as Kadenge (2009; 2010). However none of these studies, save Kadenge (2010) uses a more scientific approach to data collection such as the corpus approach to deal with the question. Apart from the studies just mentioned, there are others that made reference to the English language in Zimbabwe largely from a language planning perspective. Using the world Englishes paradigm, the current study sought to expand the exploration of ZimE from a linguistic point of view by analysing linguistic evidence to verify the mainly anecdotal and conjectural substantiation of the earlier studies, in order to settle the question of the ontological status of the variety through a more objective and verifiable criterion, thus contributing to our understanding of the variety.

5.2 Summary of findings

The objectives of this study were to explore the linguistic features characteristic of ZimE and analyse possible reasons behind their occurrence. The study intended to use the linguistic features identified to settle the variety status of ZimE. This would be achieved through a compilation of a corpus of ZimE. Overall, the use of the corpus-based methodology provided more scientific evidence for the identity, distribution and possible causes of the linguistic features attested in ZimE. Sections 5.2.1, 5.2.2 and 5.2.3 below summarise the findings of the study in terms of the research objectives that have just been outlined.

5.2.1 The morphosyntactic features of Zimbabwean English

The study showed that there are indeed a number of linguistic features that ZimE shares with other African L2 varieties of English, to varying degrees. These include: inverted word order in indirect questions, the addition of a *to*-infinitive where Standard English has a bare infinitive, the use of *like* as a focussing device, the use of the resumptive pronoun, the use of *too*, *too much*, *very much* for the qualifier *very*, as well the extension of the progressive to stative verbs. The analysis also showed that there are also other features that have been attested in other African varieties of English which did not occur at all or were very rare in the Zimbabwean variety. These include the extension of the progressive to habitual contexts, the use of *she/her* to refer to inanimate things, the use of double determiners, the use of *me* instead of *I* in co-ordinate subjects, the use of *meself/myself* instead of *I* in co-ordinate subjects, different count and mass noun distinctions resulting in use of plural where Standard English has a singular form, the use of *like* as a quotative device and use of *come*-based future ingressive markers. Finally, it emerged from the analysis that there is one feature that is hardly attested in other African varieties of English i.e. the deletion of *be* before the word *gonna* which was found to be quite pervasive. Summaries of the findings on each of these features are given in subsequent paragraphs of this section.

Among the features that ZimE shares with other varieties of English is the extension of the progressive to stative verbs, a feature that was also found to be very prevalent in Black South African English as well as in most West and East African English varieties. In this regard, ZimE was found to resemble White Zimbabwean English. The extension of the progressive aspect to stative verbs, which Van Rooy (2006; 2014) has explored extensively in the context of South Africa is the only feature under tense and aspect which was found to be fairly pervasive in the ZimE data. It emerged from the analysis that the extension of the progressive aspect to stative verbs occurred most frequently in the spoken registers of the corpus.

From the agreement category of the linguistic features emerged two features which can be safely said to be characteristic of ZimE. The first one is the deletion of the auxiliary *be* before the progressive aspect and the second one is that of deletion of the auxiliary *be* before the phrase *gonna*. As already stated, the second feature is the only one among the features attested in ZimE

which turned out to be very pervasive in the ZimE data while rare in most varieties of African English. The feature was found to be totally absent in East African Englishes including Tanzanian English, Kenyan English and Ugandan English.

Apart from the above, the use of resumptive pronouns, as conceptualised by Mesthrie (1997) which falls under the relativization category of linguistic features also emerged as a salient feature of ZimE. This finding is in agreement with the literature, which shows that this feature is also fairly prevalent in both East and West Africa as well as in South African varieties of English.

One other feature (from the complementation category) which was attested in the ZimE data is the use of the *to*-infinitive in contexts where Standard English has a bare infinitive. This is a feature that was also found to be fairly prevalent in almost all African varieties of English.

The use of *too*, *too much* or *very much* for the qualifier *very* was also found to be a salient feature of ZimE though this was seen not to be as prevalent as the features that have been discussed so far. This feature was found to be quite pervasive in the peer African varieties in general as well as South African varieties in particular though it emerged that the conceptualisation of this linguistic structure varied from researcher to researcher.

The last two features that were attested in the ZimE data belong to the discourse organisation and word order category. These are the use of inverted word order in indirect questions (which was found to be most prevalent in the spoken registers of the corpus) and the use of *like* as a focussing device though the latter feature was found not to be as prevalent as reported in other African varieties of English.

It is clear from these findings that ZimE has a lot in common particularly with Black South African English as well as East African varieties. This finding offers support to the role of cross-linguistic influence as well as a shared acquisition context as I explain in the next section, in the sense that the characteristic features of these sub-Saharan varieties of English emanate from the influence of the indigenous languages of the sub-region, which are largely Bantu. Furthermore the similarity in features comes from an acquisition context whereby the English language is acquired through education.

5.2.2 Explanations for the features evident in ZimE

The occurrence of these features has been explained through reference to various factors most of which may be related to the nature of the native language of the black English language users, the acquisition context as well as the nature of the English language itself.

As for the extension of the progressive aspect to stative verbs, it emerged that the feature was a result of inadequate mastery of the rules of Standard English and there was no evidence that this feature could be a result of cross-linguistic influence. Second language users of English in Zimbabwe simply showed a lack of close knowledge regarding the intricacies of aspect. However, it was observed that the lack of a corresponding word or concept for aspect in most Zimbabwean indigenous languages could be responsible for a blanket treatment of both dynamic and stative verbs by the second language users of English in Zimbabwe. It was also observed that the extension of the progressive aspect to stative verbs could be explained in terms of simplification strategies such as overgeneralisation.

Regarding the deletion of the *be* auxiliary before verbs in the progressive form, it emerged that ZimE differed considerably from its peer varieties in the sense that the use of the feature in ZimE is much depressed as this proved to be a feature in which the variety closely sticks to Standard English conventions. The reason for the minimal occurrence of this feature was that the English used by second language users is acquired through the education system by and large.

The reasons behind the occurrence of the deletion of the auxiliary *be* before *gonna* were found to be a most interesting case of cross-linguistic influence. It emerged that the closeness in spelling of the informal phrase *gonna* to the Shona word *gona* is a plausible explanation for the occurrence of the feature in ZimE. It was also observed that the fact that *gonna* looks like one word and in the infinitive form would give an impression that such a word would need no auxiliary before it. This was found to imply that the occurrence of this feature can be attributed to both intralingual and interlingual factors.

As for the use of the resumptive pronouns, which were found to be configured in a rather unique way, the main explanation for the occurrence of the feature was found to have something to do with cross-linguistic factors as well as pragmatic needs of spontaneous communication such as

keeping track of the subject or noun phrase as observed by Mesthrie (1997:132). It was also noted that the ZimE corpus presents new perspectives to the structural configurations and pragmatic constraints attendant upon the use of the feature.

Regarding the addition of *to* where Standard English has a bare infinitive, it emerged that the few cases where the feature occurred could be explained through reference to the fact that the second language users of English have had very little exposure to vicarious cases in which the feature is used in particular contexts. This conclusion was strengthened by the fact that the use of the *to* infinitive where Standard English would use the bare infinitive mainly occurred in student writing which in itself is a developing variety. There was little evidence of direct cross-linguistic influence for the occurrence of the feature from the data, though the overgeneralisation factor may not be ruled out.

As for the use of *too much*, *too*, *very much* for *very* qualifier, some of the reasons for the occurrence of the feature were found to be related to the overlapping meanings between *too* and *very*. The other reason was also found to be related to cross-linguistic influence whereby there are not always corresponding equivalents for *very* and *too* in the indigenous languages of Zimbabwe.

Though not very prevalent, the occurrence of inversion of word order in indirect questions was seen to be related to fossilised language learner varieties. It was also observed that the discrepancies in the grammatical packaging of indirect questions between English and the indigenous languages of Zimbabwe also contributed to the occurrence of inversion.

Lastly, the use of *like* as a focussing device which was found to be largely in keeping with Standard English conventions was argued to be a result, not of acquisition through the education system, but through interaction between the second language users and acrolectal speakers on one hand and between second language users and native users of English though admittedly to a limited extent.

However, it also emerged from the study that there are a number of features which were seen to be low-frequency features or unattested features in ZimE. These features can be said to negatively characterise ZimE. The features unsurprisingly include those that Kortmann and

Lunkenheimer (2012) have hypothesised to be typically absent from African Englishes. The features, as already mentioned, include the use of *she* for inanimate things, the use of *me/meself/myself* instead of *I* in coordinate subjects, use of plural form where Standard English uses singular form, the use of double determiners, the extension of the progressive aspect to habitual contexts, use of *come*-based ingressive markers and the use of *like* as a quotative device.

5.2.3 The status of ZimE

In terms of the variety status question of ZimE and based on corpus evidence in relation to the morphosyntactic features of ZimE, it emerged from this study that notwithstanding clear evidence of nativisation, ZimE may not be regarded as a conventional variety of English. In fact according to Schneider's (2003; 2007) Dynamic Model, ZimE may be said to be at the exonormative phase (phase 2) and the nativisation phase (phase 3) at the same time. It was observed that the indigenisation drive being encouraged by the government in most sectors of Zimbabwean life including political and economic ideologies and policies would have given the birth and development of a peculiar ZimE variety much impetus, but alas, the development of the variety seems to have fossilised because of lack of codification and acceptance as well as seemingly negative attitudes towards an indigenised English variety. It would look like the country missed a propitious opportunity to develop its own English variety just after independence in 1980, a form of Event X in Schneider's (2003) description of the birth and development of new Englishes.

5.3 Contributions of the study

Broadly speaking and against the backdrop of scanty research on ZimE, this study, which has explored the linguistic features and variety status of ZimE, has certainly reduced that research gap. In particular, the study has complemented previous studies that largely relied on impressionistic and conjectural or anecdotal evidence by building a comprehensive and so far the biggest corpus on ZimE, which can be used as a source of evidence. Though not as big as more time and more resources would have permitted, the corpus may put future explorations of the

linguistic identity of ZimE on a more concrete footing. It is also possible to explore a range of research questions on ZimE and from different perspectives using the corpus.

It should be pointed out that the data analysis in Chapter 4 of this thesis incorporated as far as possible explanations for the occurrence of the different features of the variety. These explanations helped take the treatment of ZimE as a new variety of English from a descriptive plane to a more analytical one. It is possible to glean general processes of language use in the context of language contact from the explanations, thus contributing to the theoretical frontiers of the language sciences as well as an understanding of humanity in general.

5.4 Limitations of the study and suggestions for further research

Though the findings from this corpus-based study clearly show important pointers in regard to the morphosyntactic characteristics and variety status of ZimE, due to a rather small corpus, a truly quantitative analysis of the morphosyntactic features was not possible to achieve as would be desirable. It remains important to conduct future studies that use a bigger corpus (including multimodal dimensions and biographical data of the research participants) of the variety to verify the findings. The scope of this study did not also incorporate a detailed study of the attitudes of the second language users of English in Zimbabwe though existing literature tends to point to negative attitudes. It would thus be important for future research to examine the question of attitudes towards ZimE in particular in order to expand the sociolinguistic perspective bearing on the status of the variety. Such knowledge would contribute to an understanding of the current status of ZimE and resolution of its ontological status.

Out of the 235 morphosyntactic features identified in the taxonomy proposed by Kortmann and Lunkenheimer (2012), which taxonomy was used in the present study, it turned out that only 8 features were pervasive (compared to Kenyan and Black South African English which have 14 features apiece), potentially implying that ZimE is characterised by surprisingly different morphosyntactic features or, simply by phonological features most importantly, lexicon second most importantly and morphosyntactic features only marginally.

Thirdly, the fact that no sociolinguistic information of the speakers was gathered during data collection made it impossible to resolve the variety status conclusively since, as indicated earlier this question has to be seen in a wider (sociolinguistic, cognitive and political) context.

5.5 Overall conclusion

An exploration of the linguistic features characteristic of ZimE has been hypothesised in this study to be central to the resolution of the status question of the variety. Therefore the study singled out a number of features in order to fully describe them as well as explain how they may have emanated. This exploration has been helpful in shedding light on the morphosyntactic features characteristic of ZimE and how the variety compares with similar peer varieties on one hand and with Standard English on the other. More importantly this exploration has pointed to the fact that ZimE manifests clear evidence of nativisation, though the absence of codification and other indicators of institutionalisation of the variety would preclude it from passing as a fully-fledged new variety of English. On a prognostic note, the current political and socioeconomic circumstances in Zimbabwe, despite the official national thrust towards indigenisation of all aspects of Zimbabwean life and the ever-dwindling numbers of native speakers of English, seem not to guarantee further development of the variety into the endonormative stage. ZimE seems to have clearly fossilised astride the exonormative and nativisation phases of Schneider's (2003; 2007) Dynamic Model.

References

- Aarts, J. & Meijs, W. 1984., eds. *Corpus linguistics: Recent developments in the use of computer corpora in English Language research*. Amsterdam: Radopi.
- Adolphs, S. & Carter, R. 2013. *Spoken corpus linguistics*. London: Routledge.
- Akmajian, A., Demers, R.A., Farmer, A.K. & Harnish R.M. 2010. *Linguistics: an introduction to language and communication*. Massachusetts: The MIT Press.
- Angogo, R. & Hancock, I. 1980. English in Africa: Emerging standards or diverging regionalisms? *English World-Wide*, 1(1):68-96.
- Arua, E. A. 1998. Some syntactic features of Swazi English. *World Englishes*, 17 (2):139-152.
- Arua, E. A. 2004. Botswana English: Some syntactic and lexical features. *English World-Wide*, 25(2):255-272.
- Baker, P. 2006. *Using corpora in discourse analysis*. London: Continuum.
- Bamgbose, A. 1998. Torn between the norms: Innovations in world Englishes. *World Englishes*, 17(1):1-14.
- Banjo, A. 1971. Towards a definition of standard Nigerian English. *Annales de l'Universite d'Abidjan Ser, H 1*:165-175.
- Bernsten, J. 2001. English in South Africa: Expansion and nativization in concert. *Language problems and language planning*, 25(3):219-235.
- Biber, D. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, D. 2010. Corpus-based and corpus-driven analyses of language variation and use. (In Heine, B. & Narrog H., eds. 2010. *The Oxford handbook of linguistic analysis*. Oxford: Oxford University Press, p. 159-191.)
- Biber, D., Conrad, S. & Leech, G. 2002. *Longman student grammar of spoken and written English*. Harlow: Pearson Education Limited.

- Biber, D., Conrad, S., and Reppen, R. 1998. *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. 1999. *Longman grammar of spoken and written English*. Essex: Pearson Education Limited.
- Bolton, K. 2003. *Chinese Englishes: a sociolinguistic history*. Cambridge: Cambridge University Press.
- Bolton, K. 2005. Where WE stands: approaches, issues, and debate in world Englishes. *World Englishes*, 24(1):69-83.
- Botha, Y. V. 2012. Specification in the English nominal group with reference to student writing. Potchefstroom: North-West University. (Thesis –Phd).
- Boye, K. & Harder, P. 2007. Complement-taking predicates: Usage and linguistic structure. *Studies in language*, 31(3):569-606.
- Brosnahan, L.F. 1958. English in southern Nigeria. *English studies*, 39(3):91-110.
- Bruthiaux, P. 2003. Squaring the circles: issues in modelling English worldwide. *International journal of Applied Linguistics*, 13(2):159-178.
- Buregeya, A. 2006. Grammatical features of Kenyan English and the extent of their acceptability. *English World-Wide*, 27(2):199-216.
- Buthelezi, Q. 1995. South African Black English. (In Mesthrie, R., ed. *Language and social history*. Cape Town: David Philip. p.242-250).
- Carter, R., McCarthy, M., Mark, G. & O'keeffe, A. 2016. *English language today: an A–Z of spoken and written grammar*. (Available at: <https://www.teachingenglish.org.uk/biblio/english-grammar-today-a-z-spoken-written-grammar>, accessed 17 June 2016.)
- Cheng, W. 2012. *Exploring corpus linguistics: language in action*. London: Routledge.

Chiumia, S. 2013. How many Zimbabweans live in South Africa? The numbers are unreliable. Retrieved from <http://africacheck.org>.

Chiwome, E. & Thondhlana, J. 1992. Sociolinguistics and education: A survey concerning attitude on the teaching of Shona through the medium of English. (In Herbert, R.K., ed. *Language and society: the theory and practice of sociolinguistics in Africa*. Johannesburg: Witwatersrand University Press. p.247-263).

Chivhanga, E. & Chimhenga, S. 2013. Language planning in Zimbabwe: The use of indigenous languages (Shona) as a medium of instruction in primary schools. *Journal of humanities and social science*, 12(5):58-65.

Clancy, B. 2010. Building a corpus to represent a variety of a language. (In O’Keeffe, A. & McCarthy, M., eds. *The Routledge handbook of corpus linguistics*. London: Routledge. p.80-92).

Coetzee-Van Rooy, S. & Van Rooy, B. 2005. South African English: labels, comprehensibility and status. *World Englishes*, 24(1):1-19.

Collins, P., Macalinga Borlongan, A., & Yao, X. 2014. Modality in Phillipine English. *Journal of English linguistics*, 42(1):68-88.

Criper, L. 1971. A classification of types of English in Ghana. *Journal of African languages*, 10(3):6-17.

Crystal, D. 1997. *The Cambridge encyclopedia of language* (2nd edition). Cambridge: Cambridge University Press.

Crystal, D. 2006. English worldwide. (In Hogg, R & Denison, D., eds. *A history of the English language*. Cambridge: Cambridge University Press. p.420-439).

D’Arcy, A. 2015. Quotation and advances in understanding syntactic systems. *Annual review of linguistics*, 1:43-61. Retrieved from www.annualreviews.org. (Date of access 3 June 2015.)

De Kadt, E. 2002. Gender and usage patterns of English in South African urban and rural contexts. *World Englishes*, 21(1):83-97.

- De Klerk, V. 1999. Black South African English: where to from here? *World Englishes*, 18(3):311-324.
- De Klerk, V. 2003. Towards a norm in South African English: the case for Xhosa English. *World Englishes*, 22(4):463-481.
- De Klerk, V. 2004. Procedural meanings of *well* in a corpus of Xhosa English. *Journal of pragmatics*, 37(2005):1183-1205.
- De Klerk, V. 2005. The use of *actually* in spoken Xhosa English: A corpus study. *World Englishes*, 24(3):275-288.
- De Klerk, V. & Gough, D. 2002. Black South African English. (In Mesthrie, R., ed. *Language in social history*. Cape Town: David Philip. p.365-378).
- Dube, R. & Cleghorn, E. 1999. Codeswitching in mathematics lessons in Zimbabwe. *Zimbabwe journal of educational research* (ZJER), 11(1):1-13.
- Eastwood, J. 1994. *Oxford guide to English grammar*. Oxford: Oxford University Press.
- Erling, E. 2006. Englishes/New Englishes/World Englishes. *An encyclopaedia of the arts* 4, 405-410.
- Evans, S. 2015. Testing the Dynamic Model: The evolution of the Hong Kong English lexicon (1858-2012). *Journal of English linguistics*, 43(3):175-200.
- Evison, J. 2010. What are the basics of analysing a corpus? (In O'Keeffe, A & McCarthy, M., eds. *The Routledge handbook of corpus linguistics*. London: Routledge. p.122-135).
- Fitzmaurice, S. 2010. L1 Rhodesian English. (In Schreier, D, Trudgill, P., Schneider, E.W. & Williams, J.P., eds. *The lesser-known varieties of English: an introduction*. Cambridge: Cambridge University Press. p.263-285).
- Fitzmaurice, S. 2012. White Zimbabwean English. (In Kortmann B. & Lunkenheimer, K., eds. *The Mouton world atlas of variation in English*. Berlin: Walter de Gruyter).

- Gisborne, N. 2000. Relative clauses in Hong Kong English. *World Englishes*, 19(3):357-371.
- Gough, D. 1996. Black English in South Africa. (In De Klerk, V., ed. *Focus on South Africa*. Amsterdam: John Benjamins. p.53-78).
- Gut, U. 2007. First language influence and final consonant clusters in the new Englishes of Singapore and Nigeria. *World Englishes*, 26(3):346-359.
- Gut, U. 2011. Studying structural innovations in New English varieties. (In Mukherjee, J. & Hundt M., eds. *Exploring Second Language varieties of English and learner Englishes: Bridging a paradigm gap*. Amsterdam: John Benjamins. p.101-124).
- Hans-Georg, W. & Polzenhagen, F. 2009. *World Englishes: A cognitive sociolinguistic approach*. Berlin: Walter de Gruyter.
- Hickey, R., ed. 2012. *Areal features of the Anglophone world*. Berlin: Walter de Gruyter.
- Hole, H.M. 1936. Pioneer days in Southern Rhodesia. *Journal of the Royal African Society*, 35 (138):37-47.
- Huddleston, R. & Pullum, G.K. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Hundt, M. & Vogel, K. 2011. Overuse of the progressive in ESL and learner Englishes – fact or fiction. In Mukherjee, J. & Hundt, M., eds. *Exploring the second language varieties of English and learner Englishes: Bridging a paradigm gap*. Amsterdam: John Benjamins. p.145-165).
- Hungwe, K. 2007. Language policy in Zimbabwean education: historical antecedents and contemporary issues. *Compare*, 37(2):135-149.
- James, A. 2008. New Englishes as post-geographic Englishes in lingua-franca use: genre, interdiscursivity and late modernity. *European journal of English studies*, 12(1):97-112.
- Jenkins, J. 2009. *World Englishes: a resource book for students*. London: Routledge.

- Jespersen, O. 1927. A modern English grammar on historical principles. Part III: *Syntax* (Vol II). London: George Allen & Unwin.
- Kachru, B. 1982. *The other tongue: English across cultures*. Urbana IL: University of Illinois Press.
- Kachru, B.B. 1985. Standards, codification and sociolinguistic realism: the English language in the outer circle. (In Quirk, R. & Widdoson, H., eds. *English in the world: Teaching and learning of language and literature*. Cambridge: Cambridge University Press. p.11-30).
- Kachru, B.B. 1986. *The alchemy of English: the spread, functions and models of non-native Englishes*. Oxford: Pergamon Press.
- Kachru, B.B. 1988. The sacred cows of English. *English today*, 4 (4):3-8.
- Kachru, B. 1989. Teaching world Englishes. *Indian journal of Applied Linguistics*, 15(1):85-95.
- Kachru, B.B. 1990. World Englishes and applied linguistics. *World Englishes*, 9 (1):3-20.
- Kachru, B.B. 1992. Models for non-native Englishes. (In Kachru, B. ed., *The other tongue: Englishes across cultures*. Oxford: Pergamon. p.48-74).
- Kachru, B.B. 1996. World Englishes: Agony and ecstasy. *Journal of aesthetic education*, 30(2):135-155.
- Kachru, Y. 2003. On definite reference in world Englishes. *World Englishes*, 22(4):497-510.
- Kadenge, M. 2009. African Englishes: The indigenisation of English vowels by Zimbabwean native Shona speakers. *The Journal of Pan African studies*, 3(1):156-173.
- Kadenge, M. 2010. Zimbabwean English: A sociophonological exploration. *South African Journal of African Languages* 2010(1):35-51.
- Kadenge, M. & Mudzingwa, C. 2011. Diphthong simplification through spreading: an optimality theory account. *Language Matters*, 42(1):142-161.

- Kadenge, M. & Mugari, V. 2015. The current politics of African languages in Zimbabwe. *Per Linguam: A journal for language learning*, 31(2):21-32.
- Kadenge, M. & Nkomo, D. 2011. The politics of the English language in Zimbabwe. *Language matters: Studies in the languages of Africa*, 42(2):248-263.
- Kadodo, W., Kadodo, M., Bhala, T. & Bhebe, C. 2012. The influence of teachers and students' language attitudes towards the use of Shona as medium of instruction in secondary schools. *International journal of English and literature*, 3(2):32-39.
- Kamwangamalu, N. & Moyo, T. 2003. Some characteristic features of Englishes in Lesotho, Malawi and Swaziland. *Per Linguam*, 19 (1&2):39-54.
- Kennedy, G. 1998. *An introduction to corpus linguistics*. Harlow: Addison Wesley Longman.
- Kirkpatrick, A. 2007. *World Englishes: English for international communication and English language teaching*. Cambridge: Cambridge University Press.
- Kontra, M., Phillipson, R., Skutnabb-Kangas, T., & Varady, T. 2001. *Language: a right and a resource: Approaching linguistic human rights*. Budapest: Central European University Press.
- Kortmann, B & Lunkenheimer, K., eds. 2011. *The electronic World Atlas of Varieties of English [eWAVE]*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://www.ewave-atlas.org/>, Date of access, 20 February 2016.
- Kortmann, B. & Lunkenheimer, K., eds. 2012. *The Mouton world atlas of variation in English*. Berlin: Walter de Gruyter.
- Kruger, H. & Van Rooy, B. 2016. Syntactic and pragmatic transfer effects in reported speech constructions in three contact varieties of English influenced by Afrikaans. *Language Sciences*, 56:118-131.
- Lee, D.Y.W. 2010. What corpora are available? (In O'Keeffe, A. & McCarthy, M., eds. *The Routledge handbook of corpus linguistics*. London: Routledge. p.107-121).

- Louw, W.E, and Jordan, J. 1993. Corpus of Zimbabwean English at the University of Zimbabwe computer centre. *Zambezia* XX (ii):p.131-138.
- Lowenberg, P.H. 1990. Nativisation and interlanguage in Standard English: Another look. (In Alatis, J.E., (ed). *Georgetown University Roundtable on languages and Linguistics (GURT) 1990: Linguistics, language Teaching and language Acquisition: The Interdependence of Theory, Practice and Research*. Georgetown: Georgetown University Press. p.157-179).
- Magura, B.J. 1984. Style and meaning in African English: A sociolinguistics analysis of South African and Zimbabwean English. Illinois: University of Illinois. (Thesis – PhD).
- Magura, B.J. 1985. Southern African Black English. *World Englishes*, 4(1):251-256.
- Magwa, W.2010a. Revisiting the language question in Zimbabwe: a multilingual approach to language-in education policy. *Journal of multilingual discourses* 5 (2):157-168.
- Magwa , W. 2010b. The language question in the constitution reform process in Zimbabwe: A closer look at linguistic policies outlined in the Kariba draft document. *Zimbabwe International Journal of Languages and Culture*, 1(1):7-18.
- Mair, C. 1995. Changing patterns of complementation and concomitant grammaticalisation of the verb *help* in present-day British English (In Aarts, B. & Meyer, C.F., eds. *The verb in contemporary English: Theory and description*. Cambridge: Cambridge University Press. p.258-272).
- Makalela, L. 2004. Making sense of BSAE for linguistic democracy in South Africa. *World Englishes*, 23(3):355-366.
- Makanda, A.T.P. 2009. The language question in Africa: Zimbabwe case study. Pretoria: UNISA. (Thesis-MA).
- Makoni, S.B. 1993. Is Zimbabwean English a type of new English? *African Study Monographs*, 14 (2):97-107.

- Makoni, S.B., Dube, B. & Mashiri, P. 2006. Zimbabwe colonial and postcolonial language policy and planning practices. *Current issues in language planning*, 7(4):377-414.
- Mareva, R., Kaburise, P. & Klu, E.K. 2016. Teaching ‘Standard English’ or local code-switched new Englishes in schools in Outer Circle countries: Insights from selected secondary school teachers of English in Zimbabwe. *International journal of education and multidisciplinary studies*, 4(1):113-124.
- Mareva, R. & Mapako, F.P. 2012. The prevalence and forms of code-switching and code-mixing in readers’ comments on selected news articles in *The Herald* online. *International journal of academic research in progressive education and development*, 1(4):35-43.
- Marungudzi, T. 2014. Deploying a sense of plausibility in language choice: the role of English-Shona code switching in Zimbabwean classrooms. *International journal of English and education*, 3(2):462-477.
- Marungudzi, T. 2016a. “Me too”: A sociolinguistic assessment of Zimbabwean English. *International journal of English and education*, 5(3):270-280.
- Marungudzi, T. 2016b. Towards a corpus-based study of Zimbabwean English: a state-of-the-art review and implications for further research. *International journal of English and education*, 5(2):1-13.
- Maseko, B & Dhlamini, N. 2014. Mother tongue instruction for lower primary level in Zimbabwe: A halfhearted to the development of her indigenous languages. *Journal of education and practice*, 5(6):59-65.
- McCarthy, M. & O’Keeffe, A. 2010. What are corpora and how have they evolved? (In O’Keeffe, A. & McCarthy, M., eds. *The Routledge handbook of corpus linguistics*. London: Routledge. p.3-13).
- McGinley, K. 1987. The future of English in Zimbabwe. *World Englishes*, 6(2):158-164.
- Meierkord, C. 2012. *Interactions across Englishes: linguistic choices in local and international contexts*. Cambridge: Cambridge University Press.

- Mesthrie, R. 1997. A sociolinguistic study of topicalisation phenomena in South African Black English. (In Schneider E.W., ed. *Englishes around the world: Studies in honour of Manfred Gorlach*. Amsterdam: John Benjamins. p.119-140).
- Mesthrie, R. 2002. Endogeny versus contact: Afrikaans influence in South African English. *English World-Wide*, 24(3):345-358.
- Mesthrie, R. 2003. The world Englishes paradigm and contact linguistics: Refurbishing the foundations. *World Englishes*, 22(4):449-467.
- Mesthrie, R. 2004. Black South African English: Morphology and syntax. (In Schneider, E.W. & Kortmann B., eds. *A handbook of varieties of English: a multimedia reference tool, two volumes plus CD-ROM*. Berlin: Mouton de Gruyter. p.962-973).
- Mesthrie, R. 2006. World Englishes and the multilingual history of English. *World Englishes*, 2(3/4):381-390.
- Mesthrie, R. 2008. Synopsis: Morphological and syntactic variation in Africa, South and South East Asia. (In Mesthrie, R., ed. *Varieties of English 4: Africa, South and South East Asia*. Berlin: Mouton de Gruyter. p. 624-635).
- Mesthrie, R. 2012. Black South African English. (In Kortmann, B. & Lunkenheimer, K. eds, (2012). *The Mouton world atlas of variation in English*. Berlin: Walter de Gruyter. p.493-500).
- Mesthrie, R., and Bhatt, R.M. 2008. *World Englishes: The study of linguistic varieties*. Cambridge: Cambridge University Press.
- Minow, V. 2010. *Variation in the Grammar of Black South African English* 362. Frankfurt am Main: Peter Lang.
- Mlambo, M. 2009. A survey of the language situation in Zimbabwe: Is there a Zimbabwean variety of English? If so, who speaks it? *English Today*, 25(2):18-24.
- Mollin, S. 2007. New variety or learner English: Criteria for variety status and the case of Euro-English. *English World-Wide*, 28(2):167-185.

- Mooney, A; Stilwell Peccei, J.; LaBelle, S., Engoy Henrikson, B., Eppler, E., Irwin, A., Pichler, P., Preece, S and Soden, S. 2011. *Language, society and power: An introduction*. London: Routledge.
- Mufwene, S. 2001. *The ecology of language evolution*. Cambridge: Cambridge University Press.
- Mugari, V. 2014. Code-switching in Zimbabwean urban grooves music. *Language Matters*, 45 (2):224-236, DOI: 10.1080/10228195.2014.907332.
- Mühlhäusler, P. 1986. *Pidgin and Creole Linguistics*. Oxford: Blackwell.
- Mukherjee, J. & Gries, T.S. 2009. Collostructional nativisation in new Englishes: verb construction associations in the International Corpus of English. *English World-Wide*, 30(1):27-51.
- Ndamba, G.T. 2013. *A critical review of policy on language-in-education for Africa: A case for Zimbabwe*. Pretoria: UNISA. (Thesis-PhD).
- Ndhlovu, F. 2011. No to everything British but their language: Rethinking English language and politics in Zimbabwe (2000-2008). *Language, society and culture*, 33:108-119.
- Nelson, M. 2010. Building a written corpus: What are the basics. (In O’Keeffe, A. & McCarthy, M. eds., *The Routledge handbook of corpus linguistics*. London: Routledge. p.53-65).
- Ngara, E. 1977. Bilingualism in relation to language planning and language teaching. A study of the mutual impact of Shona and English. University of London - PhD thesis.
- Ngara, E.A. 1982. *Bilingualism, language contact and language planning*. Gwelo: Mambo Press.
- Nkomo, D. 2008. Language in education and language development in Zimbabwe. *Southern African linguistics & applied language studies*, 26(3):351-362.
- Nkwe, T. & Marungudzi, T. 2015. Teachers’ perspectives on the use of English as the medium of instruction in Zimbabwean schools. *South African journal of African languages*, 35(1)43-55.
- Pennycook, A. 1998. *English and the discourses of colonialism*. London: Routledge.

- Phillipson, R. 1992. *Linguistic imperialism*. Cambridge: Cambridge University Press.
- Phillipson, R. 2000. English in the new world order. Variation on a theme of linguistic imperialism and 'world' English. (In Ricento T., ed. *Ideology, politics and language policies: focus on English*. Amsterdam: John Benjamins. p.87-106).
- Phillipson, R. 2004. English in globalisation: Three approaches. *Journal of language, identity and education*, 3(1):73-84.
- Phillipson, R. 2008. The linguistic imperialism of neoliberal empire. *Critical inquiries in language studies*, 5(1):1-43.
- Platt, J., Weber, H., & Ho, M.L. 1984. *The new Englishes*. London: Routledge.
- Platt, J.T. & Weber, H. 1980. *English in Singapore and Malaysia: status, features and function*. Kuala Lumpur: Oxford University Press.
- Prince, E.F. 1990. Syntax and discourse: A look at resumptive pronouns. *Proceedings of the sixteenth annual meeting of the Berkeley Linguistics Society*, p.482-497.
- Quirk, R. 1991. Language varieties and standard language. *JALT journal*, 11(1):14-25.
- Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. 1985. *A comprehensive grammar of the English language*. London: Longman.
- Raftopoulos, B & Mlambo, A. 2009. *Becoming Zimbabwe: a history from the pre-colonial period to 2008*. Johannesburg: Jacana Media.
- Reppen, R. 2010. Building a corpus: what are the key considerations. (In O'Keeffe, A. & McCarthy, M., eds. *The Routledge handbook of corpus linguistics*. London: Routledge. p. 31-37).
- Reppen, R. 2010. The basics of corpus design. (In Okeeffe A. & McCarthy, M., eds. *The Routledge handbook of corpus linguistics*. London: Routledge. p.31-37).

- Richards, J.C. & Schmidt, R. 2002. *Longman dictionary of language teaching and applied linguistics*. London: Pearson Education.
- Schmied, J. 1985. Attitudes towards English in Tanzania. *English World-Wide*, 6(2):237-269.
- Schmied, J. 1990. Corpus linguistics and non-native varieties of English. *World Englishes*, 9(3):255-268.
- Schmied, J. 1991. *English in Africa: An introduction*. Harlow: Longman.
- Schmied, J. 1996. English in Zimbabwe, Zambia and Malawi. (In De Klerk, V., ed. *Focus on South Africa*. Amsterdam: John Benjamins. p.301-321).
- Schneider, E.W. 2003. The dynamics of new Englishes: From identity construction to dialect birth. *Language*, 79(2):233-281.
- Schneider, E.W. 2007. *Postcolonial English: varieties around the world*. Cambridge: Cambridge University Press.
- Schneider, E.W. 2011. *English around the world: An introduction*. Cambridge: Cambridge University Press.
- Schneider, E.W. 2014. New reflections on the evolutionary dynamics of new Englishes. *World Englishes*, 33(1):9-32.
- Sergeant, P. 2012. *Exploring world Englishes: language in a global context*. London: Routledge.
- Sharma, D. 2009. Typological diversity in new Englishes. *English World-Wide*, 30(2):170-195.
- Siebers, L. 2007. Morphosyntax in Black South African English. a sociolinguistic analysis of Xhosa English. University of Duisburg-Essen. (Thesis-Ph.D).
- Siebers, L. 2013. *Morphosyntax in Black South African English: a sociolinguistic analysis of Xhosa English*. Tübingen: Narr.
- Simo Bobda, A. 2012. Cameroon English. (In Kortmann, B. & Lunkenheimer, K., eds. *The Mouton world atlas of variation in English*. Berlin: Walter de Gruyter. p.431-440).

- Sinclair, J. 1991. *Corpus, concordance and collocation*. Oxford: Oxford University Press.
- Sinclair, J. 2004. Corpus and text – Basic principles. (In Wynne, M., ed. *Developing linguistic corpora: a guide to good practice*. Oxford: Oxbow Books. p.1-16).
- Skutnabb-Kangas, T. 2000. *Linguistic genocide in education - Or worldwide diversity and human rights*. Mahwah: Lawrence Erlbaum & Associates.
- Skutnabb-Kangas, T. 2001. The globalisation of (educational) language rights. *International review of education*, 47(3-4):201-219.
- Skutnabb-Kangas, T. 2002. Language policies and education. The role of education in destroying or supporting the world's linguistic diversity. Paper presented to the World congress on language policies on 16-20 April 2002 in Barcelona, Spain.
- Skutnabb-Kangas, T. 2008. *Linguistic genocide in education - or worldwide diversity and human rights*. Delhi: Orient Blackswan.
- Skutnabb-Kangas, T. & Phillipson, R. 2010. The global politics of language: Markets, maintenance, marginalisation or murder. (In Coupland, N., ed. *The handbook of language and globalization*. West Sussex: Blackwell Publishing. p.77-100).
- Smith, N. 2005. A corpus-based investigation of recent change in the use of the progressive in British English. Lancaster: Lancaster University. (Dissertation- Phd).
- Spencer, B. 2011. International sporting events in South Africa: identity realignment, and Schneider's EVENT X. *African Identities*, 9(3):267-278.
- Szreccanyi, B. & Kortmann, B. 2007. The morphosyntax of varieties of English worldwide: a quantitative perspective. *Lingua*, 119(2009):1643-1663.
- Taiwo, R. (2012). Nigerian English. (In Kortmann, B. & Lunkenheimer, K., eds. *The Mouton world atlas of variation in English*. Berlin: Walter de Gruyter. p.411-416).
- Tan, R.S.K. & Low, E. 2010. How different are the monophthongs of Malay of Malaysian and Singapore English. *English World-Wide*, 31(2):160-189.

- Thompson, S. A. 2002. "Object complements" and conversion: towards a realistic account. *Studies in Language* 26 (1):125-164.
- Tognini Bonelli, E. 2010. Theoretical overview of the evolution of corpus linguistics. (In O'Keefe, A & McCarthy, M. eds., *The Routledge handbook of corpus linguistics*. London: Routledge. p.14-27).
- Tribble, C. 2010. What are concordances and how are they used. (In O'Keefe, A & McCarthy M., eds. *The Routledge handbook of corpus linguistics*. London: Routledge. p.167-183).
- Trudgill, P. 2004. *New-dialect formation: The inevitability of colonial Englishes*. Edinburgh: Edinburgh University Press.
- Van der Walt, J. & Van Rooy, B. 2002. Towards a norm in South African Englishes. *World Englishes*, 21 (1):113-128.
- Van Rooy, B. 2006. The extension of the progressive aspect in Black South African English. *World Englishes* 25(1):37-64.
- Van Rooy, B. 2008. A multidimensional analysis of student writing in Black South African English. *English WorldWide*, 29(3):268-305.
- Van Rooy, B. 2009. The shared core of the perfect across Englishes: A corpus-based analysis. (In Hoffman, T. & Siebers, L. eds. *World Englishes: problems, properties and prospects*. Amsterdam: John Benjamins. p.309-330).
- Van Rooy, B. 2010. Social and linguistic perspectives on variability in world Englishes. *World Englishes*, 29(1):3-20.
- Van Rooy, B. 2011. A principled distinction between error and conventionalised innovation in African Englishes. (In Mukherjee, J. & Hundt, M., eds. *Exploring Second Language varieties of English and learner Englishes: Bridging a paradigm gap*. Amsterdam: John Benjamins. p.189-208).
- Van Rooy, B. 2013. Corpus linguistic work on Black South African English. *English Today*, 29(1):10-15

- Van Rooy, B. 2014. Progressive aspect and stative verbs in outer circle varieties. *World Englishes*, 33(2):157-172.
- Van Rooy, B. & Kruger, H. 2016. The innovative progressive aspect of black South African English: The role of language proficiency and normative processes. *International journal of learner corpus research*, (2):205-228.
- Van Rooy, B. & Piotrowska, C. 2015. The development of an extended time period meaning of the progressive in Black South African English. (In Collins, P., ed. *Grammatical change in English worldwide*. Amsterdam: John Benjamins. p.465-483).
- Van Rooy, B. & Terblanche, L. 2006. A corpus-based analysis of involved aspects of student writing. *Language Matters*, 37(2):160-182.
- Van Rooy, B. & Terblanche, L. 2010. Complexity in word-formation processes in new varieties of South African English. *Southern African linguistics and applied language studies*, 28(4):407-428.
- Van Rooy, B., Terblanche, L., Haase., & Schmied, J. 2010. Register variation in African English: A multidimensional study. *English World-Wide*, 31(3):311-349.
- Van Rooy, B & Wasserman, R. 2014. Do the modals of black and white South African English converge? *Journal of English Linguistics*, 42(1):51-67.
- Verhagen, A. 2001. Subordination and discourse segmentation or: Why matrix clauses may be more dependent than complements. (In Sanders, T., Schilperoord & Spooren, W., eds. *Text representation: linguistic and psychological aspects*. Amsterdam: John Benjamins. p.337-257).
- Viriri, A. 2003. Language planning in Zimbabwe: The conservation and management of indigenous languages in Zimbabwe. Paper presented at ICOMS 14th General Assembly and Scientific symposium at Victoria Falls, Zimbabwe on the 30th of October 2003.
- Viriri, E & Viriri M. 2013. The prevalence of code-switching in secondary schools where English is the official Medium of Instruction: A case study of Buhera South District.

International journal of academic research in progressive education and development, 2(1):227-234.

Wade, R.D. 1996. An investigation of the putative restandardisation of South African English in the direction of a new English, Black South African English. Pietermaritzburg: University of Natal. (Thesis-MA).

Walt, J.L. & Rooy, B.V. 2002. Towards a norm in South African Englishes. *World Englishes*, 21(1):113-128.

Wardhaugh, R. 1998. *An introduction to sociolinguistics. (Third edition)*. Oxford: Blackwell Publishers.

Wardhaugh, R. 2006. *An introduction to sociolinguistics. (Fifth edition)*. Oxford: Blackwell Publishers.

Werner, V. 2013. Temporal adverbials and the present perfect/past tense alternation. *English World-Wide*, 34(2):202-240.

Xiao, R. 2009. Multidimensional analysis and the study of world Englishes. *World Englishes*, 28(4):421-450.

APPENDICES

Appendix 1: Sample of research participant consent form

Recording consent form

Title of thesis: The linguistic features and variety status of Zimbabwean English

Researcher

Thadeus Marungudzi

Email

tmarungudzi@gzu.ac.zw
tmarungudzi@yahoo.com or
thadeusmarungudzi@gmail.com

Cell number

0773 595 859

Address

Department of English and Performing Arts
Great Zimbabwe University
P O Box 1235
Masvingo.

Institution where researcher enrolled

North-West University
School of Languages
Vaal Triangle Campus
Vanderbijlpark
South Africa.

Promoter

Prof. Bertus Van Rooy

Please complete the following:

I, _____, hereby confirm that I have carefully read through the information page and consent form related to the above study.

I hereby confirm my willingness to participate in the study.

I also give my permission that the collected data will be used for research purposes and that no information will be published which can be used by someone else to identify myself or my organisation.

Signature of participant

Date

Appendix 2: Information sheet for research participants

Ethics information sheet for research participants

A: IMPORTANT INFORMATION

Title of thesis: The linguistic features and variety status of Zimbabwean English

Researcher

Thadeus Marungudzi

Email

tmarungudzi@yahoo.com or
tmarungudzi@cooltoad.com

Cell number

0773 595 859

Address

Department of English and Performing Arts
Great Zimbabwe University
P O Box 1235
Masvingo.

Institution where enrolled

North-West University
School of Languages
Vaal Triangle Campus
Vanderbijlpark
South Africa.

Promoter

Prof. Bertus Van Rooy

- **What is the purpose of the study?**

This study is a requirement for the completion of a PhD degree in English/Linguistics and Literary Studies. It intends to investigate the characteristics of spoken and written English used by second language users in Zimbabwe in order to establish its status in terms of being an independent variety of English.

- **How have you been selected to take part in this study?**

Second language users of English in Zimbabwe involved in different types of interaction (in English) such as spoken dialogue and spoken monologue were randomly identified and selected to take part in the research.

- **Who will have access to the information I have gathered from you?**

I, Thadeus Marungudzi, and my supervisor, Prof. Bertus Van Rooy, will have access to the recording of the interaction. However, the responsibility to keep this information safely solely lies with me.

- **Does participating in this research have any risks for you or does it have any advantages?**

There are no risks in participating in this research study. This study has the possible advantage of increasing our understanding of Zimbabwean English, language contact processes and their effect on the languages involved.

- **What are your rights?**

You are allowed to:

- withdraw from the study at any time should you wish to do so; and
- view the information that I have gathered before I start interpreting it or analysing it, should you wish to do so.

- **How is the information gathered from you going to be treated and used?**

All of the information will be handled with utmost confidentiality. Neither you nor your company's name will be shown anywhere in the discussion or findings of the research. No information will be published which can be used by someone to identify you or your organisation.

All the information will be stored in a locked filing cabinet or on a password-protected hard drive to which only I have access.

The information will only be used for the purposes of compiling a research study thesis.

B: UNDERTAKING BY PARTICIPANT

I, _____ am aware of the fact that my participation in this research study is non-compulsory and that I can withdraw from the study at any time should I wish to do so.

The researcher has explained the goals and the nature of the study to me. I understand what is expected from me and where and how the information will be used. I have the right to view the information that was gathered before it is interpreted and analysed, should I wish to do so.

All of the information will be handled with utmost confidentiality and neither me nor my organisation's name will be shown anywhere in the research report. Any discussions or findings will also not include my name or that of my organisation. No information which can be used by someone else to identify me or my organisation will be published.

I hereby give my permission to the researcher to use the data gathered for the purpose explained in this information sheet.

Signature of participant: _____ Date: _____

Signature of researcher: _____