Frequency and Variability in the Use of Modal Verbs in Zimbabwean English

Faith Chiedza Chapwanya[1] · Joanine Hester Nel[1]

Abstract

This corpus-based exploratory study examines modal verbs in Zimbabwean English (ZimE), and British English, to determine whether there are variations in use between the two varieties. Data from the ZimE corpus and the International Corpus of English Great Britain (ICE-GB) were analysed quantitatively and statistically to gain insights into how the usage patterns of modal verbs might be influenced by the usage context. Results show variations in the frequency of *can* and *will*, with the ZimE corpus registering more frequencies compared to the ICE-GB. The frequencies of *could, may, might,* and *would* were higher in the ICE-GB compared to the ZimE corpus. Overall, the ICE-GB contained more modal verbs than the ZimE corpus. In the spoken registers, results indicate that *could, may, might, shall*, and *would* occurred more in the ICE-GB, while *can* and *will* were attested more in the ZimE corpus compared to the ICE-GB. In the written registers, the ICE-GB recorded more occurrences of *could, may, might*, and *would*. A qualitative analysis was done by means of semantic interpretations of the modal verbs in different contexts. Results show the use of *can be able to* occurring as a modal phrase. The effects of language contact also lead to code-switching. Both results serve as possible motivations for variation. Furthermore, the use of *can* and *will* in past time contexts, and the use of the habitual *will* in past contexts, account for the higher overall frequencies of *can* and *will* in the ZimE corpus.

Introduction

This paper provides a comparative analysis of, firstly, the frequency of modal verbs in Zimbabwean English (ZimE), a second language (L2) English variety used by first language (L1) Shona-speakers, and British English (BrE), an L1

✉ Faith Chiedza Chapwanya
u12167160@tuks.co.za

1 Department of Afrikaans, University of Pretoria, Pretoria, South Africa

English variety, in order to, secondly, determine whether there are differences between the two varieties. ZimE was compared to BrE because BrE is the variety that provides norms in teaching and learning, government and in the print media, even though there may be influences from other English varieties on ZimE (Kadenge, 2009). This exploratory study seeks to gain insights into how the usage patterns of modal verbs might be influenced by the context, as a pragmatic analysis, in which ZimE is used as an L2, by analysing results quantitatively and statistically to see if significant differences occur between the ZimE corpus and the ICE-GB. The study also includes a qualitative analysis based on semantic interpretations of modal verbs in different contexts.

The linguistic landscape of Zimbabwe is characterised by language contact due to the fact that many languages are spoken in the country, leading to cross-linguistic influences occurring between languages (Kachru, 1992). To highlight the diversity in the languages spoken in Zimbabwe, the Constitution of Zimbabwe Amendment Act (Number 20) of 2013 acknowledges 16 official languages including English and Shona. The use of English in Zimbabwe is preferred in high domains like education, government, industry and trade, parliamentary business, and the print media, whilst indigenous languages are marginalised and spoken mainly at home (Kadenge, 2010). Despite this use of English in high domains as mentioned above, most learners have minimal to no English language exposure outside the classroom and only become bilingual due to language exposure received at school (Kadenge, 2009). This study thus utilises the term "Zimbabwean English", to, firstly, outline if there is variation in the use of different forms of English by different speakers, those being either L1- or L2-speakers of English in different contexts and to, secondly emphasize the pluricentric approach to World Englishes in the analysis of how modal verbs are used by Shona L1 speakers, who use English as an L2 in Zimbabwe.

An analysis of such variation includes a focus not only on when and to which extent (focussing on frequency), different modal verbs are used but also how they are used (thus focussing also on semantic and pragmatic interpretations). Whilst previous research has shown that variations in the frequencies of modal verbs do occur in L2 varieties of English (e.g. Collins et al., 2014; Collins, 2009a, 2009b; Deuber, 2010; Hansen, 2018; Hinkel, 1995; Kotze & Van Rooy, 2020; Leech, 2003, 2011; Rossouw & Van Rooy, 2012; Van Rooy, 2005, 2021; Van Rooy & Wasserman, 2014; Wasserman & Van Rooy, 2014), no study has examined how modal verbs are used and the different contexts and forms in which they can occur in ZimE, an area explored by the pragmatic analysis of the current study. Thus, the following research questions were posed:

1. How do the overall frequency of occurrences and variations of nine different modal verbs compare in the ZimE corpus and the ICE-GB?
2. How do these occurrences and variations differ semantically and pragmatically, for the abovementioned nine modal verbs in terms of various registers, specifically, spoken and written registers?
3. What are the possible motivations for the variations?

An Exploration of Modal Verbs: Definitions and Previous Research

Semantic and Pragmatic Features of Modal Verbs

Modal verbs (also referred to as modals or modal auxiliaries) are understood to be the primary means used to express modality in English (see Collins, 2009b; Hansen, 2018). Modality, tense and aspect are the "three kinds of information that are often encoded by verbal morphology" (Kroeger, 2005: 147). Regarding tense, present, past, and (more arguably) future are three absolute tenses whilst aspect reveals how action unfolds (Kroeger, 2005). Modality is defined as a linguistic category which forms part of linguistic meaning making, focussing on the expression of possibility and necessity, while modal verbs can further be subdivided into different categories of semantic interpretation. Furthermore, it is necessary to note that the various semantic meanings associated with the use of modals differ depending on the context in which these modal verbs are used. Collins (2009b) thus expands the notion of modality to include necessity and possibility but also ability, obligation, permission, and hypotheticality. The apparent ambiguity, which exists in the inclusion of different notions and linguistic categories associated with modality, is also included in the fact that a lack of consensus exists among scholars regarding the compilation of a complete and all-inclusive list of modal verbs. Some scholars list nine modal verbs such as *can, could, may, might, must, shall, should, will*, and *would* (e.g. Hansen, 2018; Hykes, 2000). The nine modal verbs are referred to as "central" modal verbs by Quirk et al. (1985), and were analysed for variations in this study, based on the modal verbs being the most frequently studied, enabling comparison with results from previous studies. Other researchers include *have* (*got*) *to* the list to provide for a more expansive list of ten modal verbs (e.g. Richards & Schmidt, 2010) whilst researchers who include 11 modal verbs add *ought to,* and *need* to the list of central modal verbs (e.g. Coates & Leech, 1980; Collins, 2009a).

Regarding the meanings of the different modal verbs, they are considered to be polysemous because their functions and uses in context can overlap (see Quirk et al., 1985; Collins, 2009a). Categories used to aid in the distinction between these polysemous meanings include epistemic, deontic, and dynamic modality and are utilised by Palmer (2001) and Collins (2009a) in the classification of modal verbs. Epistemic modality ideally includes the speaker's viewpoint in relation to the truthfulness of the situation and how a speaker determines the probability of truthfulness of the idea on which the statement is based, situated between weak possibility and strong possibility (Collins, 2009a). Deontic modality denotes obligation and permission and "with deontic modality the conditioning factors are external to the relevant individual" (Palmer, 2001: 9). Ability or willingness are expressed by dynamic modality (Palmer, 2001).

Quirk et al. (1985), Collins (2009a) and Wasserman and Van Rooy (2014) further categorise modals into three groups namely, (i) permission, ability and possibility (*can, could, may,* and *might*), (ii) obligation and necessity (*should* and *must*) as well as (iii) prediction and volition (*will, would* and *shall*). This

study seeks to establish whether the variations observed in the frequencies of occurrence of modals between ZimE and BrE can be accounted for based on the semantic and pragmatic shades highlighted above.

While the data analysis of the corpora included an analysis of nine different modal verbs (*can*, *could*, *may*, *might*, *must*, *shall*, *should*, *will* and *would*), the following section places emphasis on the different meanings of *can, could*, *will*, and *would* in terms of both semantic interpretations and pragmatic cues. Discussions on tense include three aspects because they assist in distinguishing the different meanings of modals as they are used in context: (i) past (hypothetical) and nonpast uses (present and future); (ii) meanings, including possibility (dynamic or epistemic ability), deontic permission, offer, requests, instructions, suggestions, prediction and volition, and habitual use; and (iii) the use of modals in terms of apodoses and protases. The above-mentioned analyses will help to account for the variations between *can* and *could* and between *will* and *would* observed in the data. Previous research has explored the different meanings of *can* and *could* as illustrated by the discussion of examples (1) to (6) below taken from Deuber (2010). The meanings include nonpast uses (which refer to the present and sometimes the future) showing possibility (dynamic or epistemic) as in example (1), where *could* expresses the likelihood of what is being referred to as the reason for the lack of *command of the English language*.

(1) Or maybe that *could* be a part of the problem why the children don't have a good command of the English language (ICE-Trinidad and Tobago (T&T) S1A-011: Deuber, 2010: 121)

The nonpast use of *can* is shown in the sentences "*can you read it*" (Deuber, 2010: 118), where the speaker is asking whether the hearer is willing or capable of reading and *could* expresses dynamic ability or capability in "*Oh right. Well could you think in your language*" where the speaker asks if the hearer is capable or willing to think in his or her language (Deuber, 2010: 118). *Can* and *could* are also associated with perception or cognitive verbs which are used to convey what is experienced through the senses as highlighted in (2).

(2) Well I think that is part of it certainly because uhm <unclear>words</ unclear> clearly influenced by those other kind of factors and uhm I mean you *could see* it in the males that . . . it is not<,> fashionable it is not the thing to achieve it's not the thing to be involved it is not the thing to do to participate (Deuber, 2010: 121)

Deontic permission which refers to duty or obligation used in nonpast and non-hypothetical context is shown in (3). Deuber (2010) suggests that this type of meaning is not attested in L1 English varieties.

(3) No it's OK you *could* sit but button up your blouse (ICE-T&T S1B-005).

With regard to past or hypothetical (imagined or suggested) uses of *can* and *could*, the two categories include past time or backshift and hypothetical meanings as exemplified in (4) and (5) respectively. The context in (4) shows that there is reference to the past as illustrated by the use of the past tense modal *could*. There is an assumption that another goal was possible if the *referee's assistant* had not ruled for *a handball* in example (5).

(4) But I mean yeah it's even at a higher level than that though because I mean Pete *could* do basic things on a computer but it wasn't enough (ICE-GB S1A-005)

(5) Soon after there *could* have been a second goal down . . . <#>But the referee's assistant called for a handball that did not appear on the replay (ICE-T&T S2B-002)

Can and *could* are also used in pragmatically specialised uses such as in offers, requests, instructions, and suggestions. For instance, in example (6), *could* is used to request the hearer to reveal the *divider*.

(6) And a divider<,> *could* you show me what's a divider<,> <#>All right could you tell me what a divider is used for now (ICE-T&T S1B-002)

Regarding *will* and *would,* the nonpast and nonhypothetical uses include three categories of meanings. The first category includes prediction (dealing with forecasting the future) and volition or willingness. The second category encompasses habitual uses of *will* and *would* involving customary, recurrent or constant habits. An example is *Oil will float on water* (Coates 1983: 178). In the preceding sentence, due to the fact that oil is less dense compared to water, it will float on top of water. The third category is the epistemic meaning which expresses "the speaker's attitude towards the factuality of the situation" (Collins, 2009a: 21). For instance, on hearing the doorbell ring, the speaker remarks *That'll be the postman* (Quirk et al., 1985: 228). The past or hypothetical (imagined or suggested) uses of *would* include the past time or backshift, habitual, and hypothetical uses. Example (7) illustrates habitual *would* where the speaker is referring to past constant habit of locking gates.

(7) That's why at least we here in La Romaine you know we *would* lock the gates at during break and lunchtime and make the students stay in the foyer (ICE-Trinidad and Tobago S1A-011, Deuber et al., 2012)

The uses of *will* and *would* in apodoses (clauses conveying consequence) of conditional sentences involves three meanings. The meanings are distinguished by the fact that the protasis or the clause that expresses the condition in a conditional sentence may be in the present tense, past tense (backshift) or past tense (see Huddleston & Pullum, 2002). As an illustration, past time reference is expressed in the protasis of the sentence *<u>If they batted first</u> they will probably win* (Huddleston & Pullum, 2002: 743). *Will* and *would* are also used in protases of conditional sentences as exemplified in (8).

(8) So, if I *would* become a person who *will* make money, I would supplement, subsidise some of their projects (ZimE corpus: Private semi-scripted dialogues 5)

The next subsection provides an overview of previous research on modals in L1 and L2 English varieties.

Previous Research on Modal Verbs

A sizeable quantity of literature on the use of modals in different varieties of English is available. The bulk of early research concentrated on L1 English varieties. For example, Leech's (2003) exploratory analysis of six different corpora representing BrE and American English (AmE) showed a notable decline in the frequency of modals and directly links to the first research question this study seeks to examine. Similar trends were observed for Canadian English by Dollinger (2008) with a focus on historical and contemporary English varieties. Millar (2009), in contrast, argued on the basis of the TIME magazine corpus, that despite a decline in the frequency of some modals, the overall picture shows an increase in the frequency of other modals. A further study by Leech (2011) maintains the assertion that overall, modals are declining in frequency. Grammaticalisation, democratisation and colloquialisation are given as the reasons for the decrease in the frequency of modals in the abovementioned studies. An overview of the literature thus shows that there are two overall trends to consider when studying the frequency of modals occurring in language use, the first pointing to an overall decline and the second to a decline with reference to only specific modals.

Some studies on modals usually compare the frequencies of modals and quasi-modals. Quasi-modals are "periphrastic modal forms, a somewhat loosely-defined grouping, formally distinguishable from, but semantically similar to the modal auxiliaries" (Collins, 2009a: 15). Although the list is not exhaustive, quasi-modals include *have to*, *have got to*, *need to*, *had better*, *be supposed to*, *be to*, *be bound to*, *be able to*, *be about to*, *be going to*, and *want to* (Collins, 2009a; Van Rooy & Wasserman, 2014). A more detailed discussion on quasi-modals falls outside the scope of this paper but can be found in Collins (2009a, 2009b) and Leech (2003, 2011). Collins (2009a) studied BrE, AmE and Australian English (AuE) and found that AmE was the variety leading in the decline of modals and the increase of quasi-modals followed by AuE and then BrE. The trend for AuE is reinforced in Collins' (2014) and Collins and Yao's (2014) studies where there are lower frequencies of modals compared to AmE. A further finding from Collins' (2009b) examination of Inner Circle varieties (BrE, AmE, AuE, and New Zealand English) and Outer Circle varieties from Philippines, Singapore, Hong Kong, India and Kenya shows that the frequency of modal verbs is lowest in AmE whilst quasi-modals are rising significantly. Regarding Outer Circle varieties, Philippine English (PhilE), Singapore and Hong Kong English recorded higher frequencies of quasi-modals compared to Kenyan and Indian English. Collins credited stylistic variation as the cause of the

higher frequency of quasi-modals in spoken registers whilst the modals' frequency was higher in the written registers.

Turning to L2 English varieties and focussing on studies linked to the second research question this study examines, Hinkel's (1995) comparison of the use of modals by L1 and L2 speakers of English in essays suggests that the use of modals of obligation and necessity (*must*, *have to*, *should*, *ought to* and *need to*) in L1 and L2 writing is influenced by culture and context. Results from Van Rooy's (2005) analysis of the expression of modality meanings in Black South African English (BSAfE) using the Tswana Leaner English (TLE) corpus, show a connection between modals and modal adjuncts. In addition, the modals *can, will* and *should* are reported to convey a stronger force in the TLE corpus whilst modal adjuncts *I think* and *maybe* are used more frequently as downtoners. With regard to modality in BSAfE, the closest and most relevant comparison to ZimE, reference is made to the use of the modal phrase *can be able to* as reflecting linguistic innovation in BSAfE (e.g. De Klerk, 2003; De Klerk & Gough, 2002; Gough, 1996; Van Rooy, 2011).

In Cameroon English, Nkemleke (2007) noted a limited use of modal verbs and contextualisation in the use of some modals, which he attributed to contact between English and Cameroonian indigenous languages. Deuber (2010), in a study of *can/ could* and *will/ would* in Trinidad English, found that Trinidad English speakers used *will* more frequently in present habitual contexts compared to BrE speakers. Deuber (2010) links this frequency to the influence of the English-based Creole on Trinidad English. The widespread use of nonpast modal verbs *will* and *can* in L2 English varieties is an interesting observation occurring in Deuber's (2010) comparison of Trinidadian data with three corpora of L1 English varieties (ICE-GB, ICE-Ireland, and ICE-New Zealand), a corpus of English as a second dialect (ICE-Jamaica), and five corpora of L2 English varieties (East Africa, Hong Kong, India, Philippines, and Singapore).

Deuber et al.'s (2012) study is pertinent since it shows variations in the use of *will* and *would* in new English varieties compared to BrE. A high frequency of the habitual *will* and the habitual *would* was attested more in non-past, non-hypothetical contexts in Trinidadian English and Bahamian English. In Jamaican English, the habitual *will* occurred less frequently compared to its frequency in Bahamian English and Trinidad English. Regarding Fiji English, the high prevalence of the habitual *would* in connection with past events was reported, whilst substantial variations between L1 English and L2 English varieties were evident. Indian English and Singapore English had higher frequencies of *will* and lower frequencies of *would.* Singapore English also shows the use of *will* with past time meaning.

PhilE, as an Outer Circle variety (e.g. Collins et al., 2014), shows a significantly slower rate of decrease in the overall frequency of *may, might, ought to, must, should* and *shall* compared to AmE and BrE. The slower decrease in the frequency of modals can be attributed to colonial lag (Trudgill, 2004) whereby PhilE has retained use of modal verbs that are used less frequently or are no longer used by its colonial parent which is AmE. The second motivation falls on the endonormativity of PhilE because the local variety is now accepted and embraced by speakers (see Schneider, 2007), as reflected by the slight changes in the frequency of *may* and *should.* In addition, the higher frequency of *shall* in PhilE is attributed to its favourable use in

administrative writing. Hansen (2018) examined variation and change in the modal verb systems of L1 and L2 varieties of English and noted similarities in the use of *must* as a main indicator of epistemic modality. Differences were observed in the use of *must* to express epistemic modality in L2 varieties of English, with Singapore English and Hong Kong English manifesting more limited use of epistemic modality than Indian English. The author noted that a substrate language's structure plays a crucial role in the way functions of *must* develop.

For Indian English, Laliberté's (2022) study of written Indian English showed an overall increase in the frequency of modals. Regarding individual modals, *can, could* and *would* increased in frequency, whilst a decrease in the frequency of *must, might* and *shall* was reported. Laliberté further observes that there is no endonormative stabilisation in the use of modals in Indian English due to non-significant variations between Indian English and its norm-providing variety, BrE.

Methodology

Data used in this study stemmed from two corpora, namely, the ZimE corpus and the ICE-GB. The ZimE corpus is composed of 356,007 words, of which 206,007 words were compiled by the researchers and 150,000 words came from Marungudzi's (2016) corpus.[1] The texts compiled by the researchers include ten private spoken dialogue samples, 45 private semi-scripted dialogue samples, ten editorial texts, 14 samples of newspaper reportage, and five business letters. For the editorials and newspaper reportage genres, online newspaper articles from *The Herald* and *The Sunday Mail* newspapers were used. Marungudzi's (2016) corpus consists of public spoken dialogues (47 texts), editorials (one text), popular writing (one text), academic writing: examination (one text), newspaper reportage (one text), business letters (three texts), social letters (two texts), public scripted spoken monologues (two texts) and creative writing (eight texts). Popular writing and academic writing genres were excluded since both had only one sample each. Therefore, 351,171 words from the ZimE corpus were used out of the 356,007 as indicated in Table 1.

Data for the 206,007 words compiled by the researchers were gathered after ethical clearance (number: GW20181012HS) was granted by the ethics committee at the University of Pretoria. Snowball, convenience and purposive sampling, the three non-probability sampling methods (Leedy et al., 2021) were used to gather data in both rural and urban areas in Zimbabwe namely Masvingo, Gweru, Mutare and Harare. Firstly, snowball sampling which "uses a small pool of initial informants to nominate other participants who meet the eligibility criteria for a study" (Morgan, 2008: 815) was utilised. Through snowball sampling, initial participants refer the researcher to potential participants to be sampled. Secondly, using a

[1] Although permission was granted to use Marungudzi's (2016) corpus by North-West University, access to only 150,000 words out of the 390,000 words was provided because Dr. Marungudzi passed away in January 2018 before handing the corpus over to the university. This meant that the data collection process had to be adapted to collect more samples than originally planned.

Table 1 Word count for text types in the ZimE corpus and the ICE-GB

Register	Text type	Words in ZimE corpus	Words in ICE-GB
Spoken	Private dialogues	131,961	185,208
	Public scripted monologues	4193	43,061
	Public dialogues	98,760	171,062
Total for spoken register		234,914	399,331
Written	Creative writing	17,890	42,646
	Editorials	26,342	20,719
	Newspaper reportage	48,939	41,539
	Social letters	5008	31,085
	Business letters	18,078	30,491
Total for written register		116,257	166,480
Grand total		351,171	565,811

convenience sampling method whereby ease of accessibility is the criterion used to select research participants (Saumure & Given, 2008), data was gathered in the Faculty of Arts at the University of Zimbabwe. Thirdly, purposive sampling was utilised because it enabled the selection of participants with particular linguistic backgrounds that fit in the study (see Leedy & Ormond, 2014). Through purposive sampling, participants who had at least 10 years of formal education were selected. Participants had diverse language and socio-cultural backgrounds and were Shona mother tongue speakers aged at least 18 years, who used English as an L2. Regarding corpus design and annotation, this study followed similar conventions as those used in the International Corpus of English (Nelson, 2002a, 2002b) and also followed by Marungudzi (2016). The International Corpus of English (1990) is a collection of corpora from more than 20 countries, of which each corpus contains one million words of different varieties of English compiled by different teams around the world (http://ice-corpora.net/ice/). The ICE-GB (1990) also used the International Corpus of English conventions, allowing for comparisons to be made between the ZimE corpus and the ICE-GB.

The second corpus utilised is Release 2 of the ICE-GB (https://www.ucl.ac.uk/english-usage/projects/ice-gp/). The reasons for choosing this corpus are: (i) that it represents English L1 use, making it possible to compare the ICE-GB with the ZimE corpus which represents L2 English use and (ii) that BrE was introduced into Zimbabwe through colonialism and has influenced ZimE (see Section "Introduction"). The ICE-GB comprises one million words, but part of the corpus totalling 565,811 words was used in this study. To enable comparison between the ZimE corpus and the ICE-GB, only samples occurring in both corpora were used. This means that the only texts used were those representing a common set of variety categories for genres in the two corpora. With respect to comparability, the ZimE corpus contains texts from the 1990s (Marungudzi, 2016), a similar period to the time when the ICE-GB was compiled. For example, radio and television stations' archived samples were the sources of spoken texts encompassing public dialogues like discussions,

interviews and public scripted monologues such as speeches, radio and television news reports. Regarding written texts, the ZimE corpus comprises of business letters, editorials and private business letters written during the same time. Table 1 shows the word count for each text category in the ZimE corpus and the ICE-GB.

A corpus-based approach was used, allowing for both quantitative and qualitative analysis of data (Lindquist & Levin, 2018). To obtain the frequencies of the modals, the ZimE corpus was analysed using Sketch Engine Tools software (www.sketchengine.eu/). The ZimE corpus was uploaded to Sketch Engine where part-of-speech-tagging was done automatically. Concordances of the modals were generated for the purpose of manually examining the key words in the contexts in which they appeared. For the ZimE corpus, the tagset for modals (MD) and the specific modal verb was searched in the wordlist tool in Sketch Engine. In this study, the negative and contracted forms of modals were included (Collins, 2009a, 2009b; Kotze & Van Rooy, 2020). For semantic interpretations of the modal verbs in different contexts in which they occurred, standard reference grammars (e.g. Huddleston & Pullum, 2002; Quirk et al., 1985) were used as the bases for analysis. In addition, some of Deuber (2010) and Deuber et al.'s (2012) classification of uses and meanings of *can* and *will* and their past tense forms *could* and *would* respectively were used to determine the reasons for the variations. In the analysis and the sorting of modal verbs during the search, the use of contextual cues such as filled and unfilled pauses and the context surrounding collocations helped in assigning the uses and meanings of modal verbs.

To ensure accuracy in the ZimE corpus, the tagset and concordance lines were manually checked and edited by utilising the Sketch Engine manual annotation tool to ensure that all occurrences represented modal verb usage. For the ICE-GB, which comes with the international corpus of English utility program (ICEUP) retrieval software, the tagset for modals verbs AUX(modal) was searched. This enabled the automatic exclusion of instances where the words were not used as modals like where *May* (referring to the month), *must* (*a must*) functioning as nouns, and *will* when it was used as a noun or a proper name.

All absolute frequencies were normalised per 100,000 words (and are presented as such in the results sections below) and the numbers were rounded off to the nearest unit. In addition, a log likelihood test (http://ucrel.lancs.ac.uk/llwizard.html) was used to determine whether the differences were statistically significant (Rayson & Garside, 2000). For the interpretation of statistical significance, four levels were used. One asterisk (*) was used to indicate that the log likelihood > 3.84 where $p < 0.05$. For $p < 0.01$, a critical value of 6.63 or higher was deemed significant and two asterisks (**) were used. The third critical value of 10.83 or higher was regarded as significant at $p < 0.001$ and three asterisks (***) were used. At level $p < 0.0001$, four asterisks (****) were used where the log likelihood value > 15.13. A plus sign appearing before the log likelihood value indicates that the higher value occurred in the ZimE corpus and a minus sign appearing before the log likelihood value indicates that the higher value occurred in the ICE-GB. In Sections "Reasons for variations between can and could" and "Reasons for variations between will and would", a log likelihood test was not performed to determine whether the variations were statistically significant because some of the frequencies of the meanings of the four

Table 2 Overall frequencies of modal verbs (normalised per 100,000 words)

Modals	ZimE corpus		ICE-GB		Log likelihood
	Absolute frequency	Normalised frequency	Absolute frequency	Normalised frequency	
Can	1585	451	1916	339	+ 70.73****
Could	449	128	1089	192	− 55.99****
May	130	37	403	71	− 46.51****
Might	105	30	459	81	− 102.66****
Must	202	58	331	59	− 0.04
Shall	80	23	115	20	+ 0.61
Should	387	110	642	113	− 0.21
Will	1669	475	2262	400	+ 28.45****
Would	758	216	2770	490	− 458.60****
Total	5365	1528	9987	1765	− 73.78****

[a]Symbols for all tables: *significant difference at $p < 0.05$, **significant difference at $p < 0.01$, ***significant difference at $p < 0.001$, ****significant difference at $p < 0.0001$

modals were very low. This was done to lower the risk of accepting false positive findings as was done in other studies with similar finding of low frequencies (see Wasserman & Van Rooy, 2014).

Results and discussion

In this section, a comparative analysis was done to check the frequencies of modals in the ZimE corpus and the ICE-GB to answer research question 1. This was done by calculating the normalised frequency and log likelihood values of the modals in each corpus and in the different registers. The section is divided into three parts. Section "Overall frequency of modal verbs" reports on the overall frequencies of modals, whilst the frequencies of modals in the spoken and written registers are provided in Section "Frequency of modal verbs across registers" (see research questions 1 and 2). Possible motivations for variations (see research question 3) between ZimE and BrE are discussed in Section "Reasons for variations between the corpora".

Overall Frequency of Modal Verbs

Table 2 displays the overall frequency of modals in the ZimE corpus and the ICE-GB. The log likelihood values are also presented.

The main quantitative difference between ZimE and BrE is the significantly lower frequency of *would* in ZimE in comparison to BrE. *Would* occurred 490 times per 100 000 words in the ICE-GB compared to 216 times in the ZimE corpus. *Can* and *will* were more frequent in the ZimE corpus (451 and 475 occurrences respectively) compared to 339 and 400 instances respectively in the ICE-GB. Variations were also observed because *could*, *may* and *might* were more frequent in the ICE-GB, with

Table 3 Frequency of modals in spoken registers (normalised per 100,000 words)

Spoken registers	ZimE corpus		ICE-GB		Log likelihood
	Absolute frequency	Normalised frequency	Absolute frequency	Normalised frequency	
Can	1254	534	1518	380	+ 77.98****
Could	315	134	722	181	− 20.31****
May	105	45	231	58	− 4.95*
Might	77	33	298	75	− 47.91****
Must	109	46	209	52	− 1.05
Shall	20	9	82	21	− 14.64***
Should	223	95	401	100	− 0.46
Will	1063	453	1402	351	+ 38.39****
Would	517	220	1981	496	− 312.48****
Total	3683	1568	6844	1714	− 19.16****

frequencies of 192, 71, and 81 respectively. In contrast, the frequencies of *could*, *may* and *might* in the ZimE corpus are 128, 37, and 30 respectively. Regarding the frequencies of *must*, *shall* and *should*, no significant differences were recorded. Overall, the ICE-GB had more modals (1765 per 100 000 words) than the ZimE corpus (1528 per 100 000 words).

Comparisons with other African English varieties show that in Cameroon English, the overall frequency of modal verbs was reported to be lower compared to BrE and AmE (Nkemleke, 2007) whilst in BSAfE, the overall frequency and the frequencies of *must* and *should* was higher (Kotze & Van Rooy, 2020; Van Rooy, 2005; Van Rooy & Wasserman, 2014).

Frequency of Modal Verbs Across Registers

In Table 3, the frequencies of modals in the spoken registers are displayed for the ZimE corpus and the ICE-GB, together with the log likelihood values. The results reveal highly significant differences in the frequency of *can* and *will* in the spoken register. *Can* occurred 534 times and *will* occurred 453 times in the ZimE corpus. In contrast, fewer occurrences of *can* and *will* were recorded in the ICE-GB (380 and 351 respectively). As specified previously in the methodology section, all the numbers reported in the discussions refer to normalised frequencies per 100 000 words. The modals *could*, *may*, *might*, *shall* and *would* were more frequent in the spoken register, in the ICE-GB compared to the ZimE corpus. Of these, the greatest difference has been found for *would*, which occurred 496 times in the ICE-GB compared to 220 times in the ZimE corpus. There were no significant variations in the frequencies of *must* and *should* between the two corpora in the spoken register.

Table 4 represents the frequency of modals in the written registers where highly significant differences are observable, with *would* recording 474 instances in the ICE-GB compared to 207 in the ZimE corpus. *Could* (220), *may* (103) and *might*

Table 4 Frequency of modals in written registers (normalised per 100,000 words)

Spoken registers	ZimE corpus		ICE-GB		Log likelihood
	Absolute frequency	Normalised frequency	Absolute frequency	Normalised frequency	
Can	331	285	398	239	+ 5.48*
Could	134	115	367	220	− 45.05****
May	25	22	172	103	− 76.73****
Might	28	24	161	97	− 61.75****
Must	93	80	122	73	+ 0.40
Shall	60	52	33	20	+ 20.63****
Should	164	141	241	145	− 0.07
Will	606	521	860	517	+ 0.03
Would	241	207	789	474	− 143.41****
Total	1682	1447	3143	1888	− 79.50****

(97) were also more frequent in the written register in the ICE-GB than in the ZimE corpus where *could* (115), *may* (22), and *might* (24) were less prevalent in the written register. There was a higher number of instances of *can* and *shall* in the written register in the ZimE corpus (285 and 52 respectively) than in the ICE-GB (239 and 20 respectively). A different trend emerges in the frequency of *must, should* and *will* in the written register because there were no significant differences.

If Tables 3 And 4 are considered, the total frequencies of modals in both registers were higher in the ICE-GB (1714 for spoken registers and 1888 for written registers) than in the ZimE corpus (1568 and 1447 respectively).

Reasons for Variations Between the Corpora

This section answers research question 3 regarding the possible motivations for the variations observed. Based on the high frequency of *can* and *will* in the ZimE corpus and on observations from previous research which also reports on the high frequency of the modals in some L2 English varieties such as Singapore English and Indian English (Deuber, 2010; Deuber et al., 2012), their uses and meanings were compared with their past tense equivalents *could* and *would* respectively.

Reasons for Variations Between *Can* and *Could*

The uses and meanings of *can* and *could* were compared in Table 5.

Table 5 shows that the non-past dynamic ability meaning of *can* was far higher in the ZimE corpus compared to the ICE-GB (288 compared to 143 respectively). In addition, *can* was also used in past time contexts in the ZimE corpus (4) compared to zero occurrence in the ICE-GB. Examples (9) and (10) indicate the past time uses of *can* in the ZimE corpus.

Table 5 Uses and meanings of *can* and *could* (normalised per 100,000 words)

Uses	Meanings	Modal verbs	ZimE corpus		ICE-GB	
			Absolute frequency	Normalised frequency	Absolute frequency	Normalised frequency
Non-past, non-hypothetical uses	Possibility (dynamic/epistemic)	*Can*	438	125	744	131
		Could	141	40	582	103
	Dynamic: ability	*Can*	1013	288	807	143
		Could	0	0	0	0
	Dynamic: perception/cognition verbs	*Can*	11	3	118	21
		Could	0	0	0	0
	Deontic: permission	*Can*	33	9	142	25
		Could	3	1	14	2
Past/hypothetical uses	Past time/backshift (all meanings)	*Can*	13	4	0	0
		Could	234	67	314	55
	Hypothetical (all meanings)	*Can*	1	0	0	0
		Could	59	17	89	16
Pragmatically specialised uses	Requests	*Can*	24	7	46	8
		Could	4	1	43	8
	Instructions	*Can*	1	0	9	2
		Could	0	0	0	0
	Suggestions	*Can*	26	7	29	5
		Could	3	1	21	4
	Offer	*Can*	2	1	5	1
		Could	0	0	0	0
Unclear/indeterminate		*Can*	23	7	16	3
		Could	5	1	26	5

Table 5 (continued)

Uses	Meanings	Modal verbs	ZimE corpus		ICE-GB	
			Absolute frequency	Normalised frequency	Absolute frequency	Normalised frequency
Total		*Can*	1585	451	1916	339
		Could	449	128	1089	192

(9)   <$H>: And Strive Masiyiwa also helped with money and so that people *can* buy clothes. (ZimE corpus: Private semi-scripted dialogue 8).

(10)   <$Q> The problem was when I was at eh Evelyn girls high, I did not score the highest marks so that I *can* join the other team who were doing science subjects. (ZimE corpus: Private semi-scripted dialogue 17)

There is reference to the past in the context of example (9) as evidenced by the use of the past tense verb *helped*. Therefore, *can* is used in the past time context where participant <$H> is narrating past events that occurred during cyclone Idai. In example (10), participant <$Q> is referring to the fact that he failed to join the other team due to not scoring high marks. The context is in the past as shown by the use of past tense verbs *was, did*, and *were*. All the 13 instances of past time *can* were attested only in private semi-scripted dialogues from different speakers. This can be attributed to the nature of the genre because participants were responding to questions and discussing past, present and future events. Since the participants use English as an L2, their language proficiency levels were varied. Although linguistic proficiency was not measured in this study, biographical and language background questionnaires provided some information regarding the levels that participants considered themselves to be at (see Chapwanya, 2022). From the questionnaires completed and given the varied language backgrounds of participants, the use of *can* in past time contexts is unlikely to be attributable to lack of proficiency in English, as most of the participants were either pursuing university education or formally employed. The study assumed that learners who had at least 10 years of formal education would have been exposed to English and could communicate in English. After 10 years of formal education, most people will have obtained their Ordinary Level certificate and will be either continuing to advanced level (Form 6) or going for vocational training and to colleges. Based on the fact that the use of *can* in past time was attested in only one genre, it is difficult to determine the extent to which the usages can be considered as features of ZimE.

In the category of dynamic: perception or cognition verbs, the ICE-GB recorded significantly higher normalised frequencies of *can* compared to the ZimE corpus (21 compared to 3 respectively). The other uses and meanings of *can* that were attested more in the ICE-GB compared to the ZimE corpus are the deontic permission and the instruction as pragmatically specialised use (25 and 2 compared to 9 and 0 respectively). No variations were observed in the use of *can* in the possibility (dynamic or epistemic), and the pragmatically specialised meanings namely requests and offer. This may be due to the fact that ZimE speakers follow BrE conventions in the use of these meanings since BrE is the norm-providing variety in schools and in government business in Zimbabwe.

With regard to *could*, the past time or backshift (all meanings) uses were more prevalent in the ZimE corpus compared to the ICE-GB (67 compared to 55). On the other hand, the possibility (dynamic or epistemic) uses and the pragmatically specialised uses namely requests and suggestions were used more in the ICE-GB compared to the ZimE corpus. There were no significant variations in the meanings of *could* for deontic permission, hypothetical (all meanings), dynamic: perception or cognition verbs, instructions and offers. In some cases, such as dynamic ability,

dynamic (perception or cognition verbs which convey what is experienced through the senses), instructions and offers, the frequencies were all zero in both corpora. Meanings of *could* with zero instances retrieved are subtypes of nonpast, non-hypothetical uses and pragmatically specialised uses. Factors, which make *could* unsuitable for these uses and which would preclude the use of *could* in any variety of English, could exist. The lack of such factors outlines a dearth in explanation of the occurrence and subsequently leads to a potential avenue for future research. As with *can,* the similarities observed may be due to the conservative nature of ZimE in the use of the above-mentioned functions of *could*. An illustration of the meaning of *could* in the past time or backshift is provided in example (11). The context in which *could* is used refers to the past as indicated by the use of past tense verbs namely *appeared, lost* and *showed*.

(11) "The party" appeared to condone the excesses of Chipangano, yet it *could* not relate this to the votes it lost at the polls as residents showed their distaste for acts done in its name. (ZimE corpus: Editorials 8)

Hyperclarity, which is intended to reduce ambiguity (Williams, 1987) may be at play in the use of *can* in the ZimE corpus. Previous research has alluded to the use of hyperclarity to reduce ambiguity and to try to be as clear and as specific as possible (e.g. De Klerk & Gough, 2002). The use of *I can say* which means "I am capable of saying something" is an example of hyperclarity. The phrase was used 182 times per one million words in the ZimE corpus compared to 12 times per one million words in the ICE-GB. Example (12) shows the use of hesitation markers *ii* and *ummm*, unfilled pauses (<,>), together with the phrase *I can say* to try and explain as clear as possible, the inspiration behind starting a reality show.

(12) <$A>: What inspired you to do this eh reality TV Completion?
      <$C> <#>Okay <,> ah <,> for me ii <,> what *I can say* is that ummm you <,> when you look at Zimbabwe <,> you've got a lot of untapped talent<,> especially in the arts industry <,> but then aah <,> they don't know have a platform <,> showcase <,> themselves. (ZimE corpus: Public dialogues 34)

Reasons for Variations Between *Will* and *Would*

The uses and meanings of *will* and *would* were analysed in order to account for the higher frequency of *will* in the ZimE corpus and *would* in the ICE-GB as shown in Table 6.

The prediction or volition uses of *will* were accounted for more in the ZimE corpus (403 compared to 351). In the non-past contexts, habitual *will* was used more frequently in the ZimE corpus (16) compared to the ICE-GB (2). This may be because of the extension of *will*, with a habitual function, in new English varieties, referred to as "habitual predictive meaning" by Quirk et al. (1985: 228). An illustration from the ZimE corpus is given in example (13).

Table 6 Uses and meanings of *will* and *would* (normalised per 100 000 words)

Uses	Meanings	Modal verbs	ZimE corpus		ICE-GB	
			Absolute frequency	Normalised frequency	Absolute frequency	Normalised frequency
Non-past, non-hypothetical uses	Prediction/volition	*Will*	1416	403	1988	351
		Would	249	85	543	96
	Habitual	*Will*	56	16	14	2
		Would	10	3	14	2
	Epistemic	*Will*	41	12	74	13
		Would	23	7	64	11
Past/hypothetical uses	Past time (except habitual)/backshift	*Will*	27	8	0	0
		Would	232	66	1186	210
	Habitual	*Will*	9	3	0	0
		Would	18	5	48	8
	Hypothetical	*Will*	3	1	9	2
		Would	52	15	534	94
Pragmatically specialised uses		*Will*	5	1	0	0
		Would	93	26	85	15
Uses in apodoses of conditional sentences	Protasis: present tense	*Will*	92	26	157	28
		Would	54	15	193	34
	Protasis: past tense	*Will*	5	1	1	0
		Would	7	2	66	12
	Protasis: past tense (backshift)	*Will*	0	0	0	0
		Would	0	0	7	1
Uses in protases of conditional sentences		*Will*	7	2	1	0
		Would	9	3	14	2

Table 6 (continued)

Uses	Meanings	Modal verbs	ZimE corpus		ICE-GB	
			Absolute frequency	Normalised frequency	Absolute frequency	Normalised frequency
Unclear/indeterminate		Will	7	2	18	3
		Would	11	3	16	3
Total		Will	1669	475	2262	400
		Would	758	216	2770	490

(13) <$Q> <#> During the week I went, I go to work from Monday to Friday. <#> I start work at eight o'clock but normally I arrive at my workplace around quarter past seven and start preparing my desk so that I can start by eight o'clock. <#> Then during my lunch hour, eh I *will* be shopping around in town. <#> Then after that, around, I get back to my workstation. <#> Then around six o'clock I *will* vacate the workplace and start looking for <indig> kombis </ indig> to go home. (ZimE corpus: Private dialogues 17)

In example (13), participant <$Q> is responding to a question about what she does during the week. The context of the discussion shows the participant describing her routine which is not in the past. Participant <$Q> starts with *I went,* then corrects herself to *I go* to describe what she usually does during the week because the context is not in the past. *Will* is used to refer to nonpast habitual meanings of shopping.

Concerning other L2 English varieties, the present habitual *will* was also more prevalent in Indian English, Singapore English and Fiji English (Deuber et al., 2012). Interestingly, habitual *will* was also used in the past contexts in ZimE as exemplified in examples (14) and (15). In the context of example (14), participant <$O> is talking about the established habit of smoking that he had in the past. In example (15), when participant <$G> was asked to talk about college, he used *will* to describe what used to happen in the past and also what was customary regarding the different types of food that was available at the college. The ZimE corpus had 3 occurrences compared to the ICE-GB, which did not have a single occurrence. Similarities can be drawn between the current study and Deterding's (2007) findings on Singapore English where the occurrence of the habitual *will* in past contexts is linked to influence from Chinese as substrate language.

(14) <$O>: Uhm so the habit was was growing within me and at times I *will* go on myself on my own to smoke. (ZimE corpus: Private semi-scripted dialogue 15)
(15) <$G>: We used to eat sadza, some *will* eat green bananas, some *will* eat cassava they only know rice. (ZimE corpus: Private dialogues 4)

Another intriguing finding of this study is that *will* was used eight times in past time contexts in two genres namely private dialogues and private semi-scripted dialogues and by different speakers in the ZimE corpus compared to the ICE-GB's zero occurrences. Comparatively, Singapore English exhibits a similar trend in the use of *will* in past time contexts (Deuber et al., 2012). The variations may be due to the fact that English is learned as an L2, mostly at school and is used minimally at home (Chapwanya, 2022). Examples (16) and (17) show past time use of *will*. Past time possibility of being robbed is referred to in example (16) where participant <$Z> is referring to an accident that happened in the past. In example (17), *will* is used with past time meaning where participant <$S> is talking about past events that occurred during cyclone Idai.

(16) <$Z>: We spend the whole night at Ziro. I was scared that we *will* be robbed. (ZimE corpus: Private semi-scripted dialogue 26)

(17) <$S>: We just went under sofas and prayed to God that we *will* be safe. Unfortunately, my neighbours, my daughter's neighbours they didn't make it. (ZimE corpus: Private dialogue 10)

The use of *will* in protases of conditional sentences, as shown in example (18) was attested more in the ZimE corpus (2) compared to the ICE-GB with zero occurrences. If results from previous studies are considered, they show that L2 varieties of English, namely, Indian English, Singapore English and Fiji English exhibit the same trend of using *will* in protases of conditional sentences (Deuber et al., 2012).

(18) <#> So, if I *will* make money, I would supplement, subsidise some of their projects. (ZimE corpus: Private semi-scripted dialogue 5)

The ICE-GB showed a considerably high frequency of *would* in the non-past prediction or volition contexts, with 96 occurrences compared to 85 in the ZimE corpus. The other uses and meanings of *would* that were attested more in the ICE-GB compared to the ZimE corpus include past time and backshift (210 compared to 66 respectively), hypothetical meanings (94 compared to 15 respectively), in three contexts in apodoses of conditional sentences namely; with a present tense verb in the protasis (34 compared to 15), and with the past tense verb in the protasis (12 compared to 2 respectively). The contexts described above account for the major quantitative difference observed in this study, which is the higher frequency of *would* in the ICE-GB compared to the ZimE corpus. Although *would* was used less in the ZimE corpus, an interesting trend was observed in the frequency of *would* for pragmatically specialised uses (26 in the ZimE corpus compared to 15 in the ICE-GB). Another interesting observation is that the collocation *I would like to* featured 15 times compared to 7 times in the ICE-GB. This variation is significant because the phrase was used in four genres namely private dialogues, public dialogues, social letters, and business letters by different participants in ZimE. Therefore, the use of *would* in the ZimE corpus may be a politeness strategy by ZimE speakers (see Huddleston & Pullum, 2002). The context in example (19) shows the use of *I would like to* as a way to soften an instruction about pick up times. In addition, the use may be an attempt at hyperclarity (Williams, 1987), a feature of L2 English varieties that was reported in previous research on BSAfE (De Klerk & Gough, 2002).

(19) *I would like to* encourage all parents who leave their children at after care to pick their children at the right time. (ZimE corpus: Business letters 1)

Can be Able to as a Modal Phrase

The high frequency of *can* in the ZimE corpus compared to the ICE-GB may be credited to its use in the modal phrase *can be able to*. Some researchers have drawn attention to the use of *can be able to* as a modal phrase in which both extrinsic possibility meaning and ability meaning are combined in BSAfE (e.g. De Klerk &

Gough, 2002: 363; De Klerk, 2003: 476; Gough, 1996: 63; Mesthrie, 2008; Van Rooy, 2011: 198). According to Van Rooy (2011: 197), in L1 English varieties, "the periphrastic expression of ability, *be able to* collocates with all modal verbs except *can* and *could*". The ICE-GB did not yield any examples of *can be able to* whilst the normalised frequency for the modal phrase is 25 times per one million words in the ZimE corpus. For a more in-depth analysis of other possible semantic interpretations, the inclusion of bigger corpora might yield different results in terms of the frequency of *can be able to* in L1 English varieties such as BrE (which is an avenue for future research). Examples (20) and (21) show the use of *can be able to.*

(20) <$B>: We've got people that uh *can be able to* help us with that. Thank you very much to uhm our listener for that. (ZimE corpus: Public dialogue 32)

(21) (…) We want to bring in academics, we want those universities to help us in formulation of laws," she said. (…) "They must be dealt with efficiently and effectively so that we *can be able to* recover the loot". (ZimE corpus: Newspaper reportage 4)

Another interesting possibility is that the expression *can be able to* seems to be a case of hyperclarity (Williams, 1987) in the sense that the meaning of possibility is communicated twice as shown in (20) and (21). For instance, the context for example (21) can mean that if the academics assist in the *formulation of laws,* then it will be possible to repossess what was looted. In this context, *can be able to* could be replaced by *will be able to*. The use of a filled pause *uh* could be considered as a delay tactic by speaker <$B> as he considered what to say next.

Can be able to occurred in both spoken and written registers, and in three different genres namely newspaper reportage, public dialogues, and private dialogues. In addition, the modal phrase was used in nine different text samples, indicating that it was not used by one participant. Although *can be able to* was used twice in newspaper reportage, which are formal texts, the examples are clearly quotations, not reportage. Therefore, it can be concluded that the modal phrase is confined to the spoken register. The frequency of *can be able to* indicates that its use is not erroneous but can be regarded as an emerging feature of ZimE. If comparisons are made with other L2 varieties of English like Xhosa English, a variety of BSAfE, Van Rooy (2011) reported that *can be able to* was attested 42 times per one million words and suggested that it is an established feature showing morphosyntactic variation in Xhosa English. In this study, the high frequency of *can* in the ZimE corpus may be linked to its use in the modal phrase *can be able to*, although to a lesser extent than BSAfE.

Language Contact: Code-Switching

Language contact between Shona and English may be linked to the overall lower frequency of modal verbs in the ZimE corpus since the overall number of words in the ZimE corpus includes both English and the Shona words inserted due to code-switching. The normalised frequency of Shona modal expressions in the ZimE

corpus is 288 per one million words. Examples of Shona modal expressions are highlighted below.

(22)   <$G>: Alright. Let's ah get into our discussion. Uh *ndichatanga* nemi baba. How did you feel pakazvarwa vana ava? (ZimE corpus: Public dialogue 26)

   Translation*: Alright. Let's ah get into our discussion. Uh <u>I will start</u> with you the father. How did you feel when these children were born?*

(23)   <$H>: *Ngavaregere* kushandisa mutemo wokungoramba zvese zvese. Kana one plus one ikanzi nditwo, ivo vanoti ndithree and vanoramba vakamira pai, pana three. *Ngavatiregererewo* vanhu ava. Zviri kuitika zvose tinodii, tinoz-viziva. (ZimE corpus: Public dialogue 25)

   Translation: <u>*They must stop*</u> *using the rule of denying everything. If one plus one is two, and they say it's three and they stick to three. These people <u>must forgive us</u>. We know what is happening.*

   In example (22), the modal expression *ndichatanga* meaning "I will start" is used to express participant <$G>'s volition or intention to begin talking to the father. In this context, a filled pause *uh* is used as a cue for code-switching. Obligation or responsibility modality is expressed in the expressions *ngavaregere* meaning "*they must stop*" and *ngavatiregererewo* meaning "*they must forgive us*" in example (23).

Conclusion

The findings of the present study showed that the largest difference between the modals in ZimE and BrE is in the quantitative difference of *would*, which occurred more in the ICE-GB. There were highly significant differences in the frequencies of *can* and *will*, with the ZimE corpus recording more occurrences for both modal verbs compared to the ICE-GB. It was also shown that, to some extent, in ZimE, *can* and *will* function as past tense modal verbs equivalent to L1 English *could* and *would* respectively. *Will* was used in past time contexts in two genres namely private dialogues and private semi-scripted dialogues while the use of habitual *will* in the past contexts was attested in creative writing and in private dialogues. Due to the limited number of genres, the features are not established in ZimE, but can be viewed as emerging. This is comparable to other L2 varieties of English such as Kenyan English (Buregeya, 2020) and Tanzanian English (Schmied, 2020) where the use of present tense forms where L1 English has past tense forms is reported to be an emerging feature. The use of *can be able to* as a modal phrase, the effects of language contact, which lead to code-switching, and hyperclarity also account for the higher overall frequency of *can* and *will* in the ZimE corpus.

   Significant differences were recorded for *would, could*, *may* and *might*, with the ICE-GB registering more occurrences than the ZimE corpus. A motivation for

these differences is the fact that BrE is an L1 variety, acquired and used both at home and at school whilst ZimE is an L2 English variety which is usually learned at school and rarely used at home (Kadenge, 2010). Therefore, the different environments where English is used as either an L1 or an L2 may result in variations in the use of modals. Regarding the overall frequencies of *must, should*, and *shall*, no variations were observed between the corpora, confirming previous research that shows a decline in their use (e.g. Collins, 2009a, 2009b; Leech, 2003, 2011). This suggests conservative behaviour by ZimE speakers since English is learned at school, and even though other English varieties may influence ZimE, BrE, a historical ancestor of ZimE, is mainly the norm-providing variety because it is used in teaching and learning, most government departments, and in the print media (see Kadenge, 2009). A comparison of ZimE and BSAfE shows an overall higher frequency of modal verbs and a marked increase in the frequencies of *must* and *should* in BSAfE (Van Rooy, 2005; Van Rooy & Wasserman, 2014; Kotze and Van Rooy 2020).

Regarding the registers, significant differences were observed in the frequency of *can* and *will* in the spoken register, where the ZimE corpus had higher frequencies compared to the ICE-GB. *Could*, *may*, *might, shall*, and *would*, were attested more in the spoken register, in the ICE-GB than in the ZimE corpus. Considering the written register, the highest difference was visible in the frequency of *would*, where the ICE-GB yielded more occurrences compared to the ZimE corpus. In addition, *could*, *may*, and *might* also showed significant differences in the written register in the ICE-GB compared to the ZimE corpus. These findings are in line with previous studies which show stylistic variation in the frequency of modal verbs in different registers (e.g. Collins, 2009a; Kotze & Van Rooy, 2020).

The link between the decline in the frequency of modals and the rise in the frequency of quasi-modals that has been reported in previous research (e.g. Collins, 2009a, 2009b; Leech, 2003) is beyond the scope of the current study and is a possible avenue for future research. Future studies may also look at the diachronic development of modals in ZimE in order to reveal possible changes in their frequencies over time. In addition, it will be interesting to determine whether there is a link between the use of modal adjuncts and the modal verbs as suggested in previous research (see Van Rooy, 2005). Another avenue to explore is the influence of Shona language on ZimE to determine if there is potential transfer between the two languages.

Declarations

Competing interests The authors state that there are no competing interests.

References

Buregeya, A. (2020). Kenyan English. In B. Kortmann, Kerstin Lunkenheimer, & K. Ehret (Eds.), *The electronic world atlas of varieties of English*. Retrieved November 14, 2023, from http://ewave-atlas.org/languages/46

Coates, J., & Leech, G. (1980). The meanings of the modals in British and American English. *York Papers in Linguistics, 8*, 23–34.

Collins, P. (2009a). *Modals and quasi-modals in English*. Rodopi.

Collins, P. (2009b). Modals and quasi-modals in Englishes. *World Englishes, 28*(3), 281–292. https://doi.org/10.1111/j.1467-971X.2009.01593.x

Collins, P. (2014). Quasi-modals and modals in Australian English fiction 1800–1999, with comparisons across British and American English. *Journal of English Linguistics, 42*(1), 7–30. https://doi.org/10.1177/0075424213512857

Collins, P., Borlongan, A. M., & Yao, X. (2014). Modality in Philippine English: A diachronic study. *Journal of English Linguistics, 42*(1), 68–88. https://doi.org/10.1177/0075424213511462

Collins, P., & Yao, X. (2014). Grammatical change in the verb phrase in Australian English: A corpus-based study. *Australian Journal of Linguistics, 34*(4), 506–523. https://doi.org/10.1080/07268602.2014.929087

Chapwanya, F. C. (2022). Investigating the use of articles, modal verbs and selected discourse markers in Zimbabwean English: A corpus-based analysis using the dynamic model. (Unpublished PhD Thesis) University of Pretoria.

De Klerk, V. (2003). Towards a norm in South African Englishes: the case for Xhosa English. *World Englishes, 22*(4), 463–481. https://doi.org/10.1111/j.1467-971X.2003.00313.x

De Klerk, V., & Gough, D. (2002). Black South African English. In R. Mesthrie (Ed.), *Language in South Africa* (pp. 356–378). Cambridge University Press.

Deterding, D. (2007). *Singapore English*. Edinburgh University Press.

Deuber, D. (2010). Modal verb usage at the interface of English and a related creole: A corpus-based study of can/could and will/would in Trinidadian English. *Journal of English Linguistics, 38*(2), 105–142. https://doi.org/10.1177/0075424209348151

Deuber, D., Biewer, C., Hackert, S., & Hilbert, M. (2012). Will and would in selected New Englishes. In M. Hundt & U. Gut (Eds.), *Mapping unity and diversity world-wide: Corpus-based studies of New Englishes* (pp. 77–102). John Benjamins.

Dollinger, S. (2008). *New-dialect formation in Canada: Evidence from the English Modal Auxiliaries*. John Benjamins.

Gough, D. (1996). Black English in South Africa. In V. De Klerk (Ed.), *Focus on South Africa* (pp. 53–77). John Benjamins.

Hansen, B. (2018). *Corpus linguistics and sociolinguistics: A study of variation and change in the modal systems of world Englishes*. Leiden: Brill Rodopi.

Hinkel, E. (1995). The use of modal verbs as a reflection of cultural values. *TESOL Quarterly, 29*, 325–343. https://doi.org/10.2307/3587627

Huddleston, R. D., & Pullum, G. K. (2002). *The Cambridge grammar of the English language*. Cambridge University Press.

Hykes, J. M. (2000). *A comparison of the use of modal verbs in research articles by professionals and non-native speaking graduate students* (Unpublished MA dissertation). Iowa State University.

International Corpus of English (1990). Retrieved December 7, 2023, from http://ice-corpora.net/ice/

International Corpus of English Great Britain (ICE-GB) (1990). Retrieved December 7, 2023, from https://www.ucl.ac.uk/english-usage/projects/ice-gb/

Kachru, B. B. (1992). Models for Non-Native Englishes. In B. B. Kachru (Ed.), *The other Tongue: English across cultures* (pp. 48–74). University of Illinois Press.

Kadenge, M. (2009). African Englishes: The indigenization of English vowels by Zimbabwean Native Shona speakers. *The Journal of Pan African Studies, 3*, 156–173.

Kadenge, M. (2010). Zimbabwean English: A Sociophonological Exploration. *South African Journal of African Languages, 30*, 35–51. https://doi.org/10.1080/02572117.2010.10587334

Kotze, H., & Van Rooy, B. (2020). Democratisation in the South African parliamentary Hansard? A study of change in modal auxiliaries. *Language Sciences, 79*, 101264. https://doi.org/10.1016/j.langsci.2019.101264

Kroeger, P. R. (2005). *Analyzing grammar: An introduction*. Cambridge University Press.

Laliberté, C. (2022). A diachronic study of modals and semi-modals in Indian English newspapers. *Journal of English Linguistics, 50*(2), 142–168. https://doi.org/10.1177/00754242221087244

Leech, G. N. (2003). Modals on the move: The English modal auxiliaries 1961–1992. In R. Facchinetti, F. R. Palmer, & M. Krug (Eds.), *Modality in contemporary English* (pp. 22–240). Mouton de Gruyter.

Leech, G. N. (2011). The modals ARE declining. Reply to Neil Millar's "Modal verbs in TIME: Frequency changes 1923–2006". *International Journal of Corpus Linguistics,* 14(2), (2009) 191–220. *International Journal of Corpus Linguistics, 16*(4), 547–564.

Leedy, P. D. & Ormond, J. E. (2014). *Practical research: Planning and design.* 10th ed. Pearson New International edition. England: Pearson Education Limited.

Leedy, P. D., Ormrod, J. E., & Johnson, L. R. (2021). *Practical research: Planning and design.* Essex: Pearson Education Limited.

Lindquist, H., & Levin, M. (2018). *Corpus linguistics and the description of English*. Edinburgh University Press.

Log Likelihood Calculator [Computer Software]. Retrieved November 15, 2019, from http://ucrel.lancs.ac.uk/llwizard.html

Marungudzi, T. (2016). Some Morphosyntactic features and variety status of Zimbabwean English: A corpus-based study. (Unpublished PhD Thesis) North-West University.

Mesthrie, R. (2008). Black South African English: Morphology and syntax. In R. Mesthrie (Ed.), *Varieties of English 4: Africa, South and South East Asia* (pp. 488–500). Mouton de Gruyter.

Millar, N. (2009). Modal verbs in TIME: Frequency changes 1923–2006. *International Journal of Corpus Linguistics, 14*(2), 191–220.

Morgan, D. L. (2008). Snowball sampling. In L. M. Given, (Ed.), *The SAGE encyclopaedia of qualitative research methods,* Vols. 1 and 2 (pp. 815–816)*.* Thousand Oaks, California: SAGE.

Nelson, G. (2002a). *International Corpus of English: Markup Manual for Spoken* Texts. Retrieved September 29, 2023, from http://ice-corpora.net/ice/index.html

Nelson, G. (2002b). *International Corpus of English: Markup Manual for Written texts*. Retrieved November 15, 2023, from http://ice-corpora.net/ICE/spoken.doc

Nkemleke, D. (2007). Frequency and use of modals in Cameroon English and application to language education. *Indian Journal of Applied Linguistics, 33*(1), 87–106.

Palmer, F. R. (2001). *Mood and modality*. Cambridge University Press.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. Longman Group Limited.

Rayson, P., & Garside, R. (2000). Comparing corpora using frequency profiling. In *Proceedings of the workshop on comparing corpora, held in conjunction with the 38th annual meeting of the association for computational linguistics* (pp. 1–6). Hong Kong.

Richards, J. C., & Schmidt, R. W. (2010). *Longman dictionary of language teaching and applied linguistics* (4th ed.). Pearson Education Limited.

Rossouw, R., & Van Rooy, B. (2012). Diachronic changes in modality in South African English. *English World-Wide, 33*(1), 1–26.

Saumure, K., & Given, L. M. (2008). Convenience sample. In L. M. Given, (Ed.), *The SAGE encyclopaedia of qualitative research methods*. Vols. 1 and 2 (pp. 124–127)*.* Thousand Oaks, California: SAGE.

Schmied, J. (2020). Tanzanian English. In B. Kortmann, K. Lunkenheimer, & K. Ehret (Eds.), *The electronic world atlas of varieties of English.* Available online at http://ewave-atlas.org/languages/45

Schneider, E. W. (2007). *Postcolonial English: Varieties around the world*. Cambridge University Press.

Sketch Engine (2003). Available online at https://www.sketchengine.eu/

Trudgill, P. (2004). *New dialect formation: The inevitability of colonial Englishes*. Edinburgh University Press.

Van Rooy, B. (2005). Expressions of modality in Black South African English. *Proceedings from Corpus Linguistics 2005*.

Van Rooy, B. (2011). A principled distinction between error and conventionalised innovation in African Englishes. In J. Mukherjee & M. Hundt (Eds.), *Exploring second language varieties of English and learner Englishes: Bridging a paradigm gap* (pp. 189–207). Benjamin Publishing Company.

Van Rooy, B. (2021). Grammatical change in South African English. *World Englishes, 40*(1), 24–37. https://doi.org/10.1111/weng.12470

Van Rooy, B., & Wasserman, R. (2014). Do the modals of black and white South African English converge? *Journal of English Linguistics, 42*, 51–67. https://doi.org/10.1177/0075424213511463

Wasserman, R., & Van Rooy, B. (2014). The development of modals of obligation and necessity in White South African English through contact with Afrikaans. *Journal of English Linguistics, 42*(1), 31–50. https://doi.org/10.1177/0075424213514588

Williams, J. (1987). Non-native varieties of English: a special case of language acquisition. *English World-Wide, 8*, 161–199. https://doi.org/10.1075/eww.8.2.02wil