Agreement in Educated Jamaican English:
A Corpus Investigation of ICE-Jamaica

Inaugural-Dissertation

zur

Erlangung der Doktorwürde
der Philologischen Fakultät
der Albert-Ludwigs-Universität
Freiburg i. Br.

vorgelegt von

Susanne Jantos
aus Räckelwitz

SS 2009
Erstgutachter: Prof. Dr. Dr. h.c. Christian Mair
Zweitgutachter: Prof. Dr. Bernd Kortmann

Vorsitzende des Promotionsausschusses
der Gemeinsamen Kommission der
Philologischen, Philosophischen und Wirtschafts-
und Verhaltenswissenschaftlichen Fakultät: Prof. Dr. Elisabeth Cheauré

Datum der Fachprüfung im Promotionsfach: 09.11.2009
Contents

1 Introduction ................................................................................................................. 1

2 Background to the Study ............................................................................................ 4
   2.1 The Jamaican Language Situation ................................................................. 4
       2.1.1 Historical Development of the Language Situation in Jamaica .......... 5
       2.1.2 New Englishes – Classification and Development ......................... 12
       2.1.3 Linguistic Variation in New Englishes and Underlying Principles ...... 20
   2.2 Linguistic Variation in Jamaica ...................................................................... 23
       2.2.1 The Creole Continuum ......................................................................... 23
       2.2.2 Morphosyntactic Characteristics of Jamaican English ..................... 27

3 Previous Research on Subject-Verb Agreement .................................................... 32
   3.1 Agreement with Lexical Verbs ..................................................................... 38
   3.2 Agreement with the Verb BE ..................................................................... 43

4 Register Variation in Subject-Verb Agreement in Educated Jamaican English 46
   4.1 Introduction ................................................................................................. 46
   4.2 Data and Methodology ............................................................................... 47
       4.2.1 Data – The International Corpus of English ..................................... 47
       4.2.2 Register-Choice for the Study of Intra-Corpus Variation ................. 49
       4.2.3 Methodology – Analysis of the Registers ......................................... 60
   4.3 Results for Spoken Jamaican English .......................................................... 63
       4.3.1 Agreement with Lexical Verbs ............................................................ 63
       4.3.2 Agreement with the Verb BE ............................................................... 83
       4.3.3 Conclusion – Subject-Verb Agreement in Spoken Jamaican English .... 92
   4.4 Results for Written Jamaican English .......................................................... 96
       4.4.1 Agreement with Lexical Verbs ............................................................ 96
       4.4.2 Agreement with the Verb BE ............................................................... 110
       4.4.3 Conclusion – Subject-Verb Agreement in Written Jamaican English ... 115
   4.5 Spoken and Written English Compared ..................................................... 118
   4.6 Conclusion .................................................................................................... 127

5 Agreement in Existential There + Be Constructions in a Cross-Variety Perspective .................................................................................................................. 130
5.1 Introduction ........................................................................................................ 130
5.2 Previous Research on Existential There + Be Constructions ....................... 131
5.3 Data and Methodology ...................................................................................... 135
  5.3.1 Data – Variety-Choice for the Study of Inter-Corpus Variation ............... 137
  5.3.2 Methodology ............................................................................................... 147
5.4 Agreement in ETBs in Native and Non-Native Varieties of English .......... 148
  5.4.1 Agreement in the Varieties ....................................................................... 148
  5.4.2 Situational Factors ................................................................................... 149
  5.4.3 Subject Number ....................................................................................... 152
  5.4.4 Tense ....................................................................................................... 154
  5.4.5 Contractedness ......................................................................................... 156
  5.4.6 Polarity ..................................................................................................... 159
5.5 Summary ......................................................................................................... 162

6   Existential Constructions ................................................................................. 165
  6.1 Expressing Existence ..................................................................................... 165
  6.2 Existential Constructions in the Substrates ................................ ................. 172
  6.3 Existential Constructions in Varieties of English ....................................... 176
  6.4 Discussion – Universal Tendencies or Language Contact? ....................... 185
  6.5 Conclusion .................................................................................................... 188

7   Concluding Remarks ......................................................................................... 190

Appendix ............................................................................................................... 202
References ............................................................................................................ 205
Zusammenfassung in deutscher Sprache ............................................................. 216
# List of Tables

1. Implicational scale – variable features of Jamaican Creole and Jamaican English ........................................ 26
2. Spoken and written text categories included ........................................................................................................ 50
3. Characteristics of the selected registers .................................................................................................................. 59
4. Distribution of lexical verbs across the spoken and written text categories ...................................................... 63
5. Distribution of agreement and non-agreement with lexical verbs in the spoken registers ......................................................... 64
6. Logistic regression for grammatical non-agreement and formality (spoken) ................................................. 81
7. Distribution of the verb *be* across the spoken and written text categories ..................................................... 83
8. Distribution of agreement and non-agreement with *be* in the spoken registers ................................................. 83
9. Distribution of agreement and non-agreement with lexical verbs in the written registers .................................................. 96
10. Logistic regression for grammatical non-agreement and formality (written) .................................................. 108
11. Distribution of agreement and non-agreement with *be* in the written registers ................................................. 110
12. Agreement and non-agreement with lexical verbs in spoken and written Jamaican English .................................................. 118
13. Agreement and non-agreement with *be* in spoken and written Jamaican English ............................................ 120
14. Complex noun phrases and proximity agreement in spoken and written Jamaican English .................................................. 125
15. Demographic characteristics and languages most frequently spoken at home in Singapore in the year 2000 .................................................. 138
16. The five scheduled languages with the highest numbers of mother tongue speakers in the 2001 census of India .................................................................................. 144
17. Distribution of agreement and non-agreement in existential *there + be* constructions .................................................. 148
18. Non-agreement with full versus contracted verb forms in category S1A .................................................. 158
19. Non-agreement by negation type .......................................................................................................................... 162
List of Figures

1. Map of creole languages of the northern Caribbean.................................................  4
2. Kachru’s concentric circles of English.................................................................  15
3. Grammatical non-agreement by subject number (spoken) .................................  65
4. Type of non-agreement across three levels of formality (spoken) .....................  67
5. Potential explanations of non-agreement involving absence of –s marking across formality levels (spoken) ..........................................................  69
6. Agreement and non-agreement with do for singular subjects in the spoken data .  76
7.  
   a. Most frequent verbs in direct conversations...............................................  78
   b. Most frequent verbs in class lessons...........................................................  79
   c. Most frequent verbs in broadcast news ......................................................  79

8. Types of non-agreement with be across three spoken registers .......................  86
9. Agreement and non-agreement in existential there + be constructions ...............  89
10. Non-agreement in ETBs by contractedness ......................................................  91
11. Grammatical non-agreement by subject number (written) ...............................  97
12. Type of non-agreement across three levels of formality (written) ....................  98
13. Potential explanations of non-agreement involving absence of –s marking across formality levels (written) ................................................... 102
14. Agreement and non-agreement with do for singular subjects in the written data 105
15.  
   a. Most frequent verbs in social letters......................................................... 106
   b. Most frequent verbs in student essays...................................................... 106
   c. Most frequent verbs in press news reports .............................................. 107

16. Types of non-agreement with be across three written registers ...................... 112
17. Types of non-agreement with lexical verbs in spoken and written Jamaican English ......................................................................................................... 121
18. Potential explanations for absent –s marking in spoken and written Jamaican English ......................................................................................................... 123
19. Agreement and non-agreement in ETBs across the whole spoken sections and in the conversational category S1A only ................................................. 150
20. Non-agreement in four varieties by text category ................................................. 151
21. Agreement and non-agreement by subject number ........................................... 152
22. Non-agreement by tense .................................................................................... 154
23. Non-agreement with full versus contracted verb forms in category S1A .......... 157
24. Non-agreement by polarity ............................................................................... 160
25. Possessive constructions with the subject pronoun they .................................. 177
26. Possessive constructions with the subject pronoun you ................................. 178
27.
   a. Constructions used to express existence in British English .................... 182
   b. Constructions used to express existence in Indian English .................... 182
   c. Constructions used to express existence in Jamaican English .............. 183
   d. Constructions used to express existence in Singapore English .......... 183
Acknowledgements

I am greatly indebted to all those who assisted me during the period of writing this book. In particular, I would like to thank Prof. Christian Mair for his support and for patiently answering all my questions at various stages of working on this book. I am grateful to Dagmar Deuber for her input and comments on several chapters and theoretical discussions of my results. Moreover, I would like to express my gratitude to Ingrid Rosenfelder for her constructive criticism and valuable advice throughout the process of writing and for the time she spent proofreading my paper. To João Menicucci I am grateful for solving various computer-related problems and I thank Seth Nyarko for his help with the statistical tests. I would also like to acknowledge the help and support received from my friends and family, especially from my parents Christine and Karl-Heinz Jantos.
1 Introduction

Jamaican English is a so-called New English (see Platt, Weber and Ho 1984) with a particularly interesting language development. The variety developed under British dominance during colonial times, but in the post-colonial era has been subject to growing influence from American English due to its closer proximity. Apart from these competing influences the language situation in Jamaica is further complicated by the presence of a substrate language, Jamaican Creole, also called ‘Patois.’ Jamaican English and Jamaican Creole are not to be seen as two strictly separate linguistic systems but rather as two varieties located at opposite ends of a standard–creole continuum with gradual and ordered transitions, ranging from an acrolect, which is the variety closest to International Standard English, to a basilect, which is farthest apart, with a number of mesolectal varieties in between (see Sand 1999: 50ff.).

Research on the Jamaican language situation in past decades has concentrated on the basilectal end of the Jamaican creole continuum (cf. Patrick 2004), while only a few studies have focused on the acrolect, the underlying assumption being that this variety was identical to standard British English, the language traditionally regarded as the norm in Jamaica due to its association with the British colonial administration. However, a number of teachers and applied linguists (cf. Christie 1989, amongst others) and later sociolinguists studying acrolectal Jamaican English have shown a tendency for it to move away from the postulated British norm (see Sand, 1999: 13-14), concluding that the emergence of a new standard, Jamaican Standard English, must be taking place (cf. Mair 2002).

It is this hypothesis which constitutes the point of departure of the present book. Jamaican English has been reported to be influenced by both British English and, due to its geographical proximity, American English (See Simo Bobda 1998a on a discussion of the influence of these varieties on World Englishes), but also by effects of Second-Language Acquisition, and vernacular universals (Chambers, e.g. 2004) applying to many or even all contact varieties of English (cf. Schneider 2003, Kortmann and Szmrecsanyi 2004, Platt, Weber and Ho. 1984, but also Mair 2003, who calls these parallel tendencies in contact Englishes ‘angloversals,’ and Simo Bobda 1998b, who labels them ‘New Englishisms’). Moreover, previous studies have pointed to a strong influence of the basilect, Jamaican Creole (cf. Mair 2002). In this complex net of influences the following question arises: Which is the factor or bundle of factors that exerts the strongest influence on educated
Jamaican English? In particular, what is the sociolinguistic status of this emerging regional standard variety of English?

These questions will be addressed in a two-pronged fashion, namely 1) through a thorough and exhaustive corpus-based analysis of subject-verb agreement in one variety (Jamaican English), and 2) through a comparative investigation of a special case of subject-verb agreement, namely agreement in existential *there + be* constructions, in Jamaican English and in four other national varieties of English represented in the International Corpus of English. The varieties chosen for this purpose are the two native reference varieties British and American English, footed on a replacement for the non-existent American ICE-corpus, as well as two non-native contact varieties, Singapore and Indian English, which were both historically British-oriented but have different kinds of substrate languages.

Apart from the type of variety there may also be social and situational conditions, such as age or gender, level of formality and characteristics of the register, which may favour or disfavour standard agreement (see Crawford 2005 for American English). In research on variation across dialects as compared to registers, genre specific grammatical variation has consistently been found to be more pronounced than regional variation (Biber et al. 1999: 20f. and Sand 2005: 74 and 167). According to Biber et al., this is because

> when speakers switch between registers, they are doing different things with language, using language for different purposes, and producing language under different circumstances. […] Regardless of any dialect differences, speakers using the same register are doing similar communicative tasks; therefore, in most basic respects the characteristic grammatical features of a register are relatively constant across speakers and dialects. (1999: 21)

To study subject-verb agreement systematically, the best choice thus seems to be a combination of an analysis of register variation and an investigation of regional, i.e. cross-variety, variation in order to gain a better understanding of the factors influencing grammatical patterns in varieties of English. In a first step the intra-variety pattern of subject-verb agreement as a central morphosyntactic feature of Jamaican English will be explored before turning to the inter-variety analysis of agreement in just one specific construction, namely *there + be* constructions, in traditional non-contact as compared to post-colonial contact varieties of English. Research on stylistic factors influencing subject-verb agreement has rarely been carried out for New Englishes, or ‘contact varieties’ (see Sand 2005), but mainly focused on native varieties of English, and the present analysis is intended to fill this gap in current research by taking a closer look at Jamaican English. The few studies that there are concentrate either on the mode of production (cf. Sand 1999
in her analysis of Jamaican radio and press texts) or on formality levels in written language, e.g. Mair, who analysed one particular written register at a time, namely Jamaican newspaper usage (1992) and student essays (2002). However, systematic investigations of such data in terms of register characteristics or effects of situational factors such as formality are still rare. For this reason, the intra-corpus analysis will focus on the level of formality and register characteristics of the texts and on differences in the mode of production, i.e. whether the texts were produced orally or in writing. For this purpose, subject-verb agreement will first be analysed in three spoken registers of educated Jamaican English that differ in their relative degrees of formality. Afterwards a similar analysis will be carried out for a comparable set of written data, before the spoken and written registers will be compared and conclusions will be drawn as to the most influential factors.

The book is structured as follows. First, a brief overview of the historical development of English in Jamaica will be given, and the sociolinguistic conditions that led to the emergence of the Jamaican standard–creole continuum characterising the present language situation in Jamaica will be discussed. Then some models for the spread and classification of varieties of English and a model accounting for the common development of New Englishes will be described. After that a short section will follow in which previous research on linguistic variation in Jamaican English will be summarized, particularly as it relates to the area of morphosyntax. Chapter three will be devoted to a theoretical overview of previous research into subject-verb agreement in different varieties of English, with a particular focus on existential \textit{there + be} constructions. In Chapter four, a description of the data and methodology applied in the investigation of subject-verb agreement is given and three spoken registers are analysed with regard to their agreement patterns and compared to the results obtained for a similar set of written data. Chapter five will focus on the inter-variety analysis of agreement in existential \textit{there + be} constructions, while chapter six explores other types of existential constructions in the same national sub-corpora of the International Corpus of English and the respective substrates. The book will end with a concluding chapter which evaluates the main findings and presents an outlook for future research.
2 Background to the Study

2.1 The Jamaican Language Situation

Jamaica is an island in the Caribbean with a population of around 2.7 million inhabitants.\(^1\) While the national and official language of Jamaica is a variety of English that has commonly been associated with British English, there is another language variety that is widely used on the island but which, in its basilectal (extreme) form, is allegedly unintelligible to speakers of Standard English (Gordon 2005). This variety, which is widely known as Jamaican Creole or ‘Patois,’ is one of the English-based creoles spoken in the Northern Caribbean, as shown on the map below. In fact, according to the Ethnologue database, Jamaican Creole rather than English “is the dominant language and gaining in prestige” in Jamaica (Gordon 2005).

![Map of creole languages of the northern Caribbean](http://www.ethnologue.com)

(Source: Gordon 2005, online at <http://www.ethnologue.com> Last updated 2009.)

**Figure 1: Map of creole languages of the northern Caribbean**

The following section will provide an overview of the historical development of the language situation in Jamaica. It will address the questions of why and how the two

\(^1\) The number given by the Statistical Institute of Jamaica for the year 2007 is 2.682 million inhabitants, while that given by the Ethnologue database (Gordon 2005, online version updated 2009) is 2.713 million.
language varieties, English and Jamaican Creole, actually came into existence on the island and how (or whether at all) they are genetically related to one another.

2.1.1 Historical Development of the Language Situation in Jamaica

The English language came to Jamaica in the 17th century and developed there in three historical main phases (cf. Mair 2003: 79ff.). The first phase began in 1655, when a British expedition of soldiers and sailors under Cromwell, who had planned to but not succeeded in capturing Santo Domingo, landed on the poorly defended Jamaican shores and seized this formerly Spanish island instead (see Holm 1986a: 8), bringing with them the English language (cf. Christie 2003: 1). Under British rule, the second half of the 17th century saw a fast increase in sugar production and in the 18th century Jamaica became one of the leading exporting nations for sugar. For this flourishing sugar industry, the British colonialists imported a large number of slaves, mainly from West Africa, to work on the sugar plantations, which led to a rapid increase of the African population in Jamaica during the 18th century (cf. Christie 2003: 8) until they far outnumbered the Europeans and by 1734 the 86,000 slaves already constituted 92% of the total population (Reinecke 1937: 288). In order to oppose this population imbalance, the colonial administration strongly encouraged immigration of indentured servants and smallholders from different regions of Britain. Most of these new arrivals were illiterate and spoke regional British dialects (cf. Christie 2003: 8). The abolition of slavery in 1838 marked the end of the first stage.

Linguistically, the policy of the slave-traders to import slaves from different African regions with different ethnic and linguistic background in order to prevent the slave groups from revolting (see Crystal 2003: 39) meant that these were not able to communicate with each other using their native languages. For this reason, the slaves employed as *lingua franca* a pidgin language, which was commonly used as a second language in Jamaica in the mid-17th century (cf. Holm 1986a: 5). There are different views on where and how this pidgin language evolved.

Some scholars suggest that the English-derived pidgin must already have developed in Africa (see Hancock 1971: 288), from where it was then transported to the New World. According to Hancock, inhabitants along the west coast had already been speaking a Portuguese pidgin due to previous contact with Portuguese-speaking sailors, and when English seamen stopped at the West African shores the coastal areas had much more
contact with English than Portuguese, so that “the Africans would have supralexified\(^2\) their Portuguese pidgin” (Hancock 1971: 289). According to this view, these Africans, after coming to the New World as slaves and being separated from other members of their original speech communities, had to continue using this pidgin to ensure communication with other slaves.

Other scholars, however, discard this theory, arguing that the slaves produced the pidgin in the first generation after their arrival in the Caribbean (see Holm 1988: 6) out of the need to communicate with fellow slaves and their masters. The beginning of this pidgin was thus the contact situation in the Caribbean between different language groups whose mother tongues were mutually unintelligible. This pidgin used as *lingua franca* in Jamaica was, however, not an arbitrary mixture of features from the languages involved but a variety characterised by a grammar that was “drastically simplified, relatively arbitrary, and devoid of inflection and derivation” (Frank 2008) and a largely English-based lexicon.

In any case, there must have been an enormous amount of linguistic variability in Jamaica as the whites spoke different regional varieties of English (cf. Sand 2002: 81) and new slaves from different linguistic backgrounds were continuously arriving. The prevailing sociolinguistic conditions on the plantations, such as the continuing need of the slaves to communicate with each other and with their overseers, led to the rapid expansion of this pidgin, as it was used for an ever-increasing number of communicative functions. As new generations of children learned this variety as their mother tongue, a “radically restructured form of English emerged over the next generations” (Sand 2002: 81).\(^3\) This English-

---

\(^2\) Hancock opts for ‘supralexification’ as “a better term [than ‘relexification’] for the lexical change involved in the adaptation of pidgins” (1971: 288), since it implies “lexicon-building rather than lexicon-replacement” (p. 288). However, there seems to be a contradiction here as he explicitly speaks of “the process of discarding Portuguese derived items” (p. 288) from the pidgin in West African coastal areas. Therefore, ‘relexification’ appears to be the more appropriate term.

\(^3\) The ongoing controversy around the development of creole languages is reflected in the wealth of literature that centers around this topic, among them Mufwene, e.g. 1986, 1999 and 2008; Muysken and Smith 1986, Lalla and D’Costa 1990, Rickford and Romaine 1999, Neumann-Holzschuh and Schneider 2000, and McWorther 2005, to mention but a few.

The views on creole genesis and particularly on possible reasons for similarities in creole languages around the world can be essentially subsumed under three main headings. The first group of scholars defend the *substratist position*, arguing that the New World creoles are “relexified forms of African languages” (Frank 2008) which derive much of their vocabulary from the respective superstrate language, e.g. English in the case of Jamaican Creole, but whose syntax can be traced to the substrate languages in Africa (cf. Frank 2008, Alleyne 1986 on Jamaica and Holm 1986b on substrate diffusion in creoles around the world).

The second view, the *superstratist position*, sees creoles as the result of “approximations of approximations of nonstandard dialects of European languages” (Frank 2008, and the work by Robert Chaudenson, e.g. 2001).

The third view of creole languages, the *Bioprogram Hypothesis* (cf. Bickerton 1981) emphasizes that the many similarities of creoles around the world cannot be explained satisfactorily by superstrate or substrate
lexifier creole called Jamaican Creole or ‘Patois’ soon became the identifying language of
the local community. Mair characterizes this creole, a form of which still exists in Jamaica
alongside English today (cf. Christie 2003: 1) as a "Kontaktsprache mit stark
westafrikanisch geprägter Phonologie, dominant englischem Vokabular und einer
neuartigen Grammatik des Kreoltyps, die durch raschen strukturellen Ausbau eines
vorgängigen Pidgin und seiner Übernahme als Muttersprache entstand" (Mair 2003: 79). The
occasional use of the creole by members of the white population was frowned upon
and considered inappropriate (cf. Mair 2003), and this sense of social stigmatization was to
remain intact also in the next developmental stage.

The second phase in the development of the language situation lasted from the abolition of
slavery in 1838 until 1962, when Jamaica gained political independence from Britain (cf.
Mair 2003). During that time, the colony lost its economic significance for the British
Empire and slowly developed socially as well as culturally (see Mair 2003). English was
spread through the education system, so that increasing fractions of the population became
proficient in both English and the creole. The use of language was no longer exclusively
socially determined, with whites speaking English and all others using only the creole, but
the distinction was of a more functional nature. English was the prestigious form
associated with the British colonialists, and used for formal and official functions as well
as in writing, while Jamaican Creole was the socially stigmatized variety, which, however,
was used for private and informal every-day communication by the whole population (cf.
Mair 2003). Thus, the sociolinguistic situation was characterized by the functional
differentiation of codes, i.e. what has come to be called ‘diglossia’ (Mair 2003: 80).

The French term ‘diglossie’ was coined by the Greek linguist Psycharis as a translation of
the Greek term for ‘bilingualism’ (δι = ‘two’ and glossa = ‘language’) (cf. Sand 1999: 39)
and was introduced into sociolinguistics by Charles Ferguson in his article “Diglossia”
(1959) to denote a kind of bilingualism of two (or more) closely related language varieties,
of which one has high social prestige, the so-called ‘high’ variety (H), and at least one has
low prestige, the ‘low’ variety (L), which is socially stigmatized (cf. Ferguson 1959).
Ferguson defines diglossia as

influence and must therefore be caused by universals of the human mind, the so-called bioprogram (cf. Frank
2008).
More recently, however, scholars have come to the conclusion that, in the creation of creole languages “the
universalist and substrate hypothesis complement one another” (as the title of an article published in 1986 by
the esteemed linguist Salikoko Mufwene suggests). Even nowadays these issues are still being hotly
disputed, but a further discussion is beyond the scope of the present book.

4 A detailed description of the population developments in Jamaica before and during colonization and the
early stages in the development of Jamaican Creole is given by Lalla and D’Costa (1990: 6ff.)
a relatively stable language situation in which, in addition to the primary dialects of the language (which may include a standard or regional standards), there is a very divergent, highly codified (often grammatically more complex) superposed variety, the vehicle of a large and respected body of written literature, either of an earlier period or in another speech community, which is learned largely by formal education and is used for most written and formal spoken purposes but is not used by any sector of the community for ordinary conversation. (Ferguson 1959: 336)

In the Jamaican case of diglossia, British English was the high prestige variety (H) and Jamaican Creole the low prestige variety (L). What Sand emphasises as the most important point is that the varieties are “in complementary distribution” (1999: 40), i.e. H is used in formal or written domains such as church sermons, political speeches, news broadcasts, education, folk literature, newspapers, etc. (see Sand 1999), in which L is not appropriate and vice versa. For instance, Ferguson (1959) describes that the use of the high variety in ordinary conversation “is not acceptable in diglossic communities” (Sand 1999: 40). However, in situations of “conquest diglossia,” as in Jamaica, where the external standard variety that functions as H was imposed by the British colonisers rather than consciously chosen, the L varieties commonly remained invisible, i.e. “they were not even deemed to exist as languages” (Devonish 2003: 160). This means that only English was regarded as real and the creole was not recognized as a variety in its own right, but rather as ‘broken English’ or ‘bad English.’

Early in this second developmental stage, both spoken and written language usage in Jamaica was diglossic. However, as Trudgill (2002) describes, language contact usually leads to the varieties influencing each other. More precisely, “[i]f a creole comes into renewed or intensified contact with its source language and the sociolinguistic conditions are right, it may begin to change in the direction of this source language” (Trudgill 2002: 70), the process of change being “decreolisation” (see Holm 1988). Due to written language being highly codified and standardized, such change usually begins in the spoken domains. This is exactly what happened in Jamaica, where the creole speakers were

---

5 Another case of diglossia is, for example, German-speaking Switzerland (cf. Devonish 2003). In contrast to Jamaica, however, the external H variety used for formal spoken and written communication in German-speaking Switzerland, i.e. High German, was not imposed on the country but is the result of a “historically pragmatic decision” (Devonish 2003: 158) taken by the Swiss themselves. Therefore, Switzerland is clearly a case of “non-conquest diglossia” (Devonish 2003: 175).

6 See Devonish (2003) for a characterization of the diglossic situation in Jamaica and the invisibility of Jamaican Creole. The later point is especially emphasized by Frank (2008), who reports that creole languages in general are often disrespected and considered “corrupted and inferior forms of a “standard” language like English or French” and that “[e]ven today, it is not uncommon to hear someone say that the creole ‘is not a real language.’” (Frank 2008)

7 See footnote 10, however, for some critical comments on ‘decreolization.’
motivated to acquire the more prestigious Standard English (cf. Holm 1988: 52). However, not only was spoken creole influenced by spoken English, but also vice versa. In this way, a language situation developed in Jamaica which can be described as a continuum\(^8\) characterized by “a cline of varieties ranging from a variety of the source language […] at the top of the social scale to increasingly un-decreolised varieties of Jamaican Creole at the bottom” (Trudgill 2002: 70). There is thus an acrolect at the upper end, namely the variety of English associated with the colonial British norm, a basilect at the other end of the continuum, in this case Jamaican Creole, and a number of intermediate forms, the so-called mesolects. In the second phase of language development, this continuum developed slowly for spoken language use, while in writing the functional diglossia was fully intact.

The third and last main stage began in 1962, when Jamaica gained political independence from Britain, and has lasted until today (cf. Mair 2003). From 1962 onwards British influence on Jamaica has been steadily decreasing. Instead, the United States has become increasingly influential, first because of its important role as a world power (cf. Holm 1986a: 17), and second because it is geographically much closer to Jamaica than Britain. Therefore, linguistically, the two native models British and American English constitute competing influences on Jamaican English (Holm 1986a: 17).

Not only is British influence growing weaker but the nature of the spoken continuum has also been slowly changing. Sand shows that the acrolect has slowly been moving away from the colonial British norm (Sand 1999: 13-14), thus indicating that a new regional standard variety of English must be evolving under the influence of the creole substrate (cf. Mair 2002, 2003). Moreover, it is nowadays difficult to observe a pure basilectal creole in Jamaica. In fact, Bailey (1966), who studied the basilect of JC, found that not even older rural speakers could be observed to use pure basilectal creole (see Görlach 2002: 32).\(^9\)

---

\(^8\) Sand (1999: 50) states that Reinecke and Tikimasa in 1934 were the first linguists who used the term ‘creole continuum’ in their attempt to account for the large amount of linguistic variation in Hawaii, but it was as early as 1914 that Schuchardt described a sociolinguistic situation in which such a continuum existed, without, however, using the term. He states: “Wo das NE [Neger Englisch] im Schatten des europäischen Englisch fortlebt, bilden sich zwischen beiden eine Menge von Zwischenstufen; es wird aus dem NE schliesslich ein Englisch der Neger, das sich von dem der Weissen nur in Aussprache, Tonfall, Redewendung unterscheidet und nicht viel mehr als etwa das Deutsche der Juden (nicht das Jüdisch-deutsche) von dem der anderen” (Schuchardt 1914: viii).

\(^9\) Since this pure creole is nowhere to be found, Bailey’s (1966) characterization of Jamaican Creole was heavily criticized by DeCamp as being “an abstract ideal type, a composite of all non-standard features, a combination which is actually spoken by few if any Jamaicans” (1971: 35). However, the other pole of the continuum, Standard Jamaican English, is also an abstract ideal, “a composite of all standard-features” (Devonish 2003: 164), as the varieties actually produced are only more or less close approximations of that standard and “even a highly educated speaker cannot be expected to use Standard Caribbean/Jamaican English to the total exclusion of creole” (Devonish 2003: 164) in all social situations. Yet these abstractions are valuable as the alleged extreme points of the creole-continuum to which ‘actual usage’ can be compared.
In spoken language use, private and public, it is becoming increasingly hard to distinguish between English and creole (cf. Mair 2003), as both varieties show a tendency to move towards the other and transitions between the mesolectal varieties of the continuum are gradual.\(^\text{10}\)

DeCamp characterizes the linguistic variability in Jamaica as follows:

The varieties of Jamaican English themselves differ to the point of unintelligibility; but some Jamaican English is mutually intelligible with standard English. […] Further, in Jamaica there is no sharp cleavage between Creole and standard. Rather there is a linguistic continuum, a continuous spectrum of speech varieties ranging from the ‘bush talk’ or ‘broken language’ of Quashie to the educated standard of Philip Sherlock and Norman Manley. (DeCamp 1971: 350)

As Jamaican speakers’ language proficiencies cover a whole span of varieties of the continuum, the breadth of which depends on their education and “the breadth of [their] social contacts” (DeCamp 1971: 350), they can “adjust [their] speech upward or downward for some distance on it” (Holm 1988: 56), depending on the particular situational circumstances.\(^\text{11}\) That means that different lects are chosen for stylistic or rhetoric purposes, often within single speech events, e.g. to mark topic transitions, emotional states or jocular language, but also to emphasize their identity as Jamaicans.\(^\text{12}\) As Hinrichs summarizes, “[i]n speech in Jamaica, the situation can be described as ‘codeswitching as

\(^{10}\) This allegedly indicates that the creole is already partially decreolised and can therefore be called a “post-creole” (Trudgill 2002: 70, cf. DeCamp 1971) in the terminology of Hall’s (1962, 1966) ‘life-cycle theory’. After the stages of contact-induced pidginization, stabilisation, expansion and creolisation, the creole continuum developed with the lexifier language at the top and the creole at the bottom of the scale (cf. Sand 1999: 51). According to proponents of this theory, Jamaican Creole has now proceeded to the last stage of the life-cycle, the decreolisation phase, in which the creole changes towards its lexifier language (see Sand 1999). This ‘decreolisation hypothesis’ (DeCamp 1971) has been controversially discussed and heavily criticized in the literature, for instance by Patrick (1999: 15-19) and Mufwene (1999). Mufwene argues that synchronic variation in the creole continuum was “treated as symptomatic of decreolisation” (1999: 158). In other words, the diachronic aspect inherent in the decreolisation hypothesis has not been sufficiently supported by diachronic evidence (Mufwene 1999: 160f., 168). He also notes that the assumption of decreolisation “is not consistent with the social history of the regions” (Mufwene 1999: 162) which would rather justify the hypothesis that creoles in the colonies “have developed in the direction of basilectalization, becoming structurally more and more different from their lexifiers” (Mufwene 1999: 162, see also, among others, Bickerton 1988, Chaudenson, e.g. 1979 and 1992). Lalla and D’Costa (1990: 98) also argue against a decreolization of Jamaican Creole during the past two centuries. Mufwene further notes that the term ‘post-creole’ is problematic, as creole languages are defined by sociohistorical and political factors rather than their structural features, so that the principle “once a creole, always a creole” (1999: 159) should apply “regardless of whatever changes it may have undergone” (Mufwene 1999: 159f.). Following Mufwene, the present book will avoid the term ‘post-creole’ and speak of the ‘creole continuum’ instead.

\(^{11}\) It is such stylistic variation that is the focus of Deuber’s (2009) study. See also section 2.2, in which the appropriateness of the creole continuum model for synchronic variation in Jamaica is discussed and some of Deuber’s findings are summarized.

\(^{12}\) See Hinrichs (2006) for a detailed discussion of codeswitching in Jamaica and the diaspora, particularly with respect to the construction of identity.
the unmarked choice”\textsuperscript{13} [...] for a steadily increasing number of domains, including informal conversation among peers” (Hinrichs 2006: 139).

In written usage, the diglossia of late colonialism is still present (cf. Mair 2003: 81), with Jamaican English being used for most formal and public purposes and Jamaican Creole being employed occasionally for rhetoric or humorous effects. There is no such thing as a creole continuum for written Jamaican language use. Even in informal computer-mediated communication (CMC) via email and Internet discussion forums, which closely resembles informal spoken communication in many respects (see Hinrichs 2006 for a very detailed discussion of the peculiarities of this rather new text type), English is consistently used for most purposes, and thus constitutes ‘the unmarked choice’ (Hinrichs 2006: 139), while the choice of creole in written language is ‘marked’, “because the cognitive cost of writing Creole is greater” (Hinrichs 2006: 139), and because Jamaicans “are ultimately more familiar with the conventions and rules of written English” (Hinrichs 2006: 139) than written creole. For this reason, written creole is mainly used for particular rhetoric effects or as explicit symbols of Jamaican identity.

The questions that arise here are the following: Are these different sociolinguistic conditions, i.e. the creole continuum in the spoken domains versus diglossia in written language, reflected in the pattern of subject-verb agreement? More precisely, is spoken language more prone to grammatical non-agreement than written language due to the more extensive mixing of English with creole features? Could it be that differences in the level of formality are the stronger factor determining (non-) agreement, accounting for considerably more non-agreement in informal communication than in formal usage, both in spoken and written domains? These questions will be addressed in chapter 4 of this book, whose first part will be devoted to stylistic and situational factors, providing separate analyses of spoken and written data, while the second part will focus on the mode of production, i.e. spoken or written.

Before turning to linguistic variation and thoroughly investigating subject-verb agreement as a central morphosyntactic feature of English and its specific characteristics in Jamaica, some models of the spread of English and the development of New Englishes around the world will be briefly discussed. The reason why such a discussion is crucial in understanding the development of Jamaican English is that it may enable us to trace parallel developments in different postcolonial varieties of English that have developed

\textsuperscript{13} Hinrichs here refers to the terminology of the Markedness Model by Myers-Scotton (1998).
under comparable or totally dissimilar socio-historical conditions. This will later help to interpret the results of the corpus-based study of subject-verb agreement in existential \textit{there + be} constructions (chapter 5) and to draw conclusions as to whether the characteristics of Jamaican English are most likely caused by influence from native languages (chapter 6), such as the former colonial British norm and the more recent point of orientation American English (cf. Mair 2003), or whether the particularly Jamaican sociolinguistic situation, namely the presence of Jamaican Creole as a substrate language, accounts for most of the variation. Alternatively, it may be possible to decide whether Jamaican English is as it is because of typical developments in postcolonial Englishes arising independently of the sociolinguistic situation due to universal tendencies of language change in all vernacular varieties of English, i.e. so-called angloversals (Mair 2003: 84).

### 2.1.2 New Englishes – Classification and Development

In the research tradition focusing on varieties of English around the world some terminological issues have been controversially discussed. One concerns the most suitable label for varieties that developed under similar sociohistorical circumstances in countries around the world. Such post-colonial Englishes have usually been called ‘New Englishes’ (cf. for example Pride 1982, Platt, Weber and Ho 1984, and Foley 1988) or ‘World Englishes,’ although both terms are not quite satisfactory. According to Kachru (1983), “the ‘New English’ of India was actually older than English in Australia, which is not generally considered ‘New’” (Mesthrie and Bhatt 2008: 3). Similarly, the term ‘World Englishes’ is problematic, as it “runs the risk of being over-general, since British English is not generally studied within this paradigm” (Mesthrie and Bhatt 2008: 3). The authors note that, similarly to the term ‘World Music’ which is used mainly for ‘non-Western’ music, the label ‘World English’ covers ‘non-Western’ varieties of English. Sand (2005) gives a detailed account on the various classifications of Englishes in the literature and opts for ‘contact varieties’ as the most suitable term.

Despite the above-mentioned terminological considerations the term ‘New Englishes’ will be used for the three post-colonial varieties investigated in this book.\textsuperscript{14} This is because they

\textsuperscript{14} The varieties Jamaican English, Singapore English and Indian English are here classified as New Englishes. British and American English are not New Englishes as they developed in completely different historical settings, i.e. they are not post-colonial contact and shift varieties in the sense understood here.
clearly fulfil the following four criteria developed by Platt, Weber and Ho (1984: 2f.). According to the authors, a variety is to be considered a ‘New English,’ if the following holds true:

1. It has developed through the education system. This means that it has been taught as a subject and, in many cases, also used as a medium of instruction in regions where languages other than English were the main languages. […].
2. It has developed in an area where a native variety of English was not the language spoken by most of the population. For various reasons […] pidgin and creole languages are not considered to be native varieties of English.
3. It is used for a range of functions among those who speak or white it in the region where it is used. This means that the new variety is used for at least some purposes such as: in letter writing, in the writing of literature, in parliament, in communication between the government and the people, in the media and sometimes for spoken communication between friends and in the family. […]
4. It has become ‘localized’ or ‘nativized’ by adopting some language features of its own, such as sounds, intonation patterns, sentence structures, words, expressions. Usually it has also developed some different rules for using language in communication. (Platt, Weber and Ho 1984: 2f.)

Furthermore, Platt, Weber and Ho (1984: 9) distinguish three types of New Englishes according to their historical background. The first type includes varieties such as Indian English, Kenyan English or Singapore English. These developed in settings where one or more local languages are present and a non-English language is used for wider communication (e.g. Hindi, Swahili, Bazaar Malay, etc.). The second type, including varieties such as Ghanaian English and Nigerian English, involves the presence of one or more local languages and an English-based pidgin used for wider communication in some areas. The third type includes, for instance, Caribbean English and involves the presence of an English-based creole. All post-colonial Englishes under investigation in the present study, i.e. Jamaican English (as a representative of Caribbean English), Indian English and Singapore English, are explicitly identified as New Englishes by Platt, Weber and Ho (1984).

Another controversial issue in the literature is the distinction between native and non-native speakers, which is increasingly recognised as problematic by researchers of World Englishes. Those persons are traditionally regarded as native speakers who have learned a given language from birth as their first language (cf. Mesthrie and Bhatt 2008: 36), while those who learn it as a second language at a later stage, at least “some time after being initiated into [their] native language” (Mesthrie and Bhatt 2008: 36), are considered to be non-native speakers. However, several scholars have argued that “such definitions of a
native speaker seem to be premised on the norms of monolingual societies, whereas in fact the world is largely multilingual” (Mesthrie and Bhatt 2008: 36). Particularly in post-colonial contexts, where different language groups are in constant and intense contact, monolingualism is the marked case and English is the first language only for a small minority of speakers. In the present study, a distinction will be made that is based on the acquisitional context in which English is learned rather than the order of acquisition. That is, a native speaker is someone who acquires English in the home without formal instruction, whereas someone who acquires English through formal education at school is a ‘non-native’ speaker. While this distinction is relevant, the present study focuses less on individual speakers and more on varieties of English. Of more immediate importance is thus the theoretical representation and classification of Englishes.

It was in the late 1980s that three scholars in the field of World Englishes presented their models of the spread of English—almost simultaneously—to the linguistic community. These models were designed to facilitate the categorization of varieties of English according to their sociohistorical development. Tom McArthur was the first of the three linguists, introducing his ‘Circle of World English,’ which was to become widely known as the ‘wheel model,’ in *ET* in July 1987, followed by Manfred Görlach, who presented a similar model at a conference in 1988 (cf. Mesthrie 2008: 29f.; See also McArthur 1987 and 1998: 97 and Görlach 1990: 42 and 2002: 100). Both McArthur’s and Görlach’s models represent the spread of English in the shape of a wheel, locating International English15 in the centre and arranging national and regional standard varieties, ‘semi-standards’ or ‘standardizing’ varieties, and ‘non-standard’ or ‘subvarieties’ in several layers around the hub of the wheel, from core to periphery. Pidgins and creoles are located at the fringe (cf. also Sand 2005: 7f.).

Braj Kachru’s ‘Concentric Circles model of English’ was the next to be published, but in fact it is this model which has come to be recognized and accepted “as the model with the best ‘fit’ for English as she has been spreading” (Mesthrie 2008: 30). In contrast to McArthur’s and Görlach’s ‘wheel,’ Kachru’s model, which is reproduced below as Figure 2, consists of three concentric circles. His classification of varieties rests on the historical background, i.e. the types of spread, how the variety of English is normally acquired, the range of functional domains in which it occurs and the societal penetration of the language

15 Görlach’s and McArthur’s models differ in that the ‘International English’ in Görlach’s model is not to be understood as a standard variety (Görlach 1990: 42) while McArthur’s ‘World Standard English’ clearly refers to an international standard variety of English (cf. Sand 2005: 7, footnote 8).
(cf. Figure 2 in Mesthrie 2008: 31). As Sand (2005: 4) observes, Kachru’s model is essentially based on a categorization of varieties into ENL, ESL and EFL countries as proposed by Moag (1982). Sand (2005: 3) summarizes Moag’s classification as follows:

ENL (English as a native language): English is an official language or is used in the domains of administration and government, is spoken as L1 by a majority of the population, is used in all domains and shows considerable internal variation (regional, social, stylistic, register, etc.)

ESL (English as a second language): English is officially recognized, used in a large number of intranational domains (e.g. administration, education, media, etc.), is spoken as L2 by a large or at least influential part of the population, shows some internal variation

EFL (English as a foreign language): English is spoken by a small percentage of the population, acquired mainly in a classroom setting, used for very specific purposes only (tertiary education, international communication, etc.)

EBL (English as a basal language): this category comprises “a small group of societies where English is the mother tongue of a minority group of a larger populace whose native tongue, often Spanish, is clearly the dominant language of the society as a whole” (Moag 1982: 11f.)

The three circles in Kachru’s model roughly correspond to Moag’s categorization of Englishes into ENL, ESL and EFL varieties.

\[\text{Figure 2: Kachru's concentric circles of English}\]
Kachru’s ‘Inner Circle’ includes the major native (ENL) varieties USA, UK, Canada, Australia and New Zealand, in which English is spoken as a mother tongue (or L1) by the majority of the population. The ‘Outer Circle’ comprises varieties that have developed in colonial settings, i.e. the ‘New Englishes’ discussed above, where English now has the status of a second language (ESL), is spoken as an L2 by a large part of the population, and is used in various internal domains, such as administration, government, education and media. Examples of outer circle countries are the post-colonial varieties Singapore English and Indian English. Last but not least, the ‘Expanding Circle’ consists of countries such as China, Nepal or Japan, where English is learned as a foreign language (EFL) and used in a limited number of domains.

While Kachru’s model is considered “a powerful visual representation of the spread of English, both geographically and historically” (Sand 2005: 6), it nevertheless has some limitations. First, Kachru (1992: 3) admits that he is unable to properly categorize varieties such as South African English or Jamaican English within this framework, as they are not typical cases of ESL varieties. In the Jamaican case, this is due to the fact that Jamaican English and Jamaican Creole are not clearly identifiable as distinct linguistic systems but in fact there is a continuum of linguistic varieties ranging from basilectal Jamaican Creole on the lower end to acrolectal Jamaican English on the upper end, with an indefinite number of gradual ordered transitions, the so-called mesolects, in between. Even those Jamaicans speaking lower mesolectal or basilectal creole consider themselves English-speakers, however, and therefore the distinction between L1 and L2, native or non-native, mother tongue or second language, is problematic in Jamaica. The second problem is that no Western European country makes an appearance in the Expanding Circle, although English is regularly used in some European countries in domains of education, tourism and international communication (Mesthrie 2008: 31). Sand mentions a third problem, namely that “Kachru’s visualization appears to imply a central role for the inner circle and a more peripheral role for the outer and expanding circles, despite the fact that he advocates the linguistic rights of the speakers in the outer and expanding circle (e.g. Kachru 1982: 35-41; 1986)” (Sand 2005: 6).

The present book focuses particularly on varieties from the outer circle, and includes

Note that Jamaican English is not mentioned separately in Kachru’s model, but Kachru does include it, albeit indirectly, as one of the Caribbean countries in the expanding circle. Cf. section 2.2.1 for a description of the linguistic systems in Jamaica.

Schneider (2003: 237 footnote 6) describes some of the problems arising in the classification of varieties of English in Kachru’s and McArthur’s models. See also Mesthrie (2008: 27ff), who discusses several different attempts to categorize Englishes and mentions some further research developments focusing on these issues.
Jamaican English on the basis of it being a post-colonial variety of English. What this means is that the categorization adopted in this book is based on the shared historical background of these varieties rather than the questionable status as ENL, ESL or EFL, and thus follows Schneider (2003), who points to a common development of post-colonial varieties. Schneider’s ‘Dynamic model of the evolution of New Englishes’ rests upon three theoretical frameworks: 1) ‘Identity Theory’ (cf. Gumperz 1982, Jenkins 1996, Woodward 1997, Eckert 2000, and Schneider 2000), a field which regards the emergence of New Englishes as the result of “sociopolitically-driven identity reconstructions” (Schneider 2003: 271); 2) Language Contact Theory (cf. Neumann-Holzschuh and Schneider 2000, Mufwene 2001), according to which “[t]he evolution of New Englishes is regarded as a special type of contact-induced language change, closely related to the genesis of pidgin and creole languages, from which it cannot be distinguished in principle” (Schneider 2003: 271); and 3) Accommodation Theory (cf. Giles 1984, Trudgill 1986), a research tradition which postulates that the evolution of New Englishes is the result of linguistic convergence, caused by the tendency of “cooperative speakers [to] approximate each other’s speech forms” (Schneider 2003: 271).

As Schneider argues, “despite all obvious dissimilarities, a fundamentally uniform developmental process, shaped by consistent sociolinguistic and language-contact conditions, has operated in the individual instances of rerooting the English language in another territory” (2003: 233). He suggests that this developmental process consists in five “characteristic stages of identity construction” (p. 233), in the course of which the STL strand, i.e. incoming Anglophone settlers, and the IDG strand, i.e. the indigenous population, increasingly interact and communicate with one another and, as a result, “a new hybrid identity of the two groups comes into existence which manifests itself in the formation of a new variety of English along the evolutionary line” (Mukherjee and Gries 2009: 30). Schneider’s (2003: 244ff.) five consecutive stages and the varying influence of the STL and IDG strands on the socio-political and linguistic development (cf. Sand 2005: 11) can be summarized as follows:

Phase 1 – Foundation: In this initial phase, large numbers of English speakers come to a previously non-English-speaking country for trade or permanent settlement and transport the English language to this new (colonial) territory. The STL strand, in which the various regional dialects of the Anglophone settlers come into close contact, is characterized by koinéization, dialect levelling, ‘focusing’ and simplification (Schneider 2003: 244). There is only restricted contact between settlers and the indigenous population, but some
indigenous people interact with the immigrants on a regular basis, and this regular interaction results in marginal bilingualism developing in the IDG strand. In this phase, the influence of indigenous languages on English is very small and restricted mainly to names for places (p. 245).

Phase 2 – Exonormative stabilization: English is now a regular means of communication in the colony which is slowly stabilizing politically under (normally) British dominance. However, the linguistic norms and standards are still determined externally. That is, “written and spoken British English as used by educated speakers, is accepted as a linguistic standard of reference” (p. 245) and the Anglophone settlers consider themselves ‘outposts of Britain’ to which they culturally belong. Nevertheless, the STL strand begins to adopt local vocabulary and gradually incorporate it into indigenous English usage, borrowing, for example, lexemes from indigenous languages or coining new words to designate local flora and fauna and later cultural objects. The IDG strand, on the other hand, begins to expand and bilingualism spreads among the indigenous population. This is because proficiency in English is increasingly valued positively as being a social advantage, “opening roads to higher status or specific commercial options” (p. 246). As both the STL and IDG strands slowly start to modify and expand their traditional identities, the process of structural nativization begins and “the earliest structural features typical of local usage emerge, if only slowly” (p. 246).

Phase 3 – Nativization: According to Schneider, this is the central and most important stage in the development of the New English, in which

both parties involved realize that something fundamental has been changing for good: traditional realities, identities, and socio-political alignments are discerned as no longer conforming to a changed reality, and the potentially painful process of gradually replacing them with something different, a new identity reflecting a changed reality, combining the old and the new, is in full swing. This process has the immediate linguistic consequences, for the drastically increased range of communication between the parties involved now makes language use a major practical issue and an expression of new identity. (Schneider 2003: 247)

In this stage, the STL and IDG strands become closely intertwined as the members of the STL strand group dissociate themselves from the mother country as the source of political power and linguistic and cultural norms and move toward gradual independence, and the indigenous people undergo a process of linguistic and cultural assimilation and large-scale second language acquisition (cf. Schumann 1978). As English becomes an integral part of the local linguistic repertoire, some “local linguistic idiosyncracies” (Schneider 2003: 248)
develop under the influence of substrate effects and interlanguage usage, including heavy lexical borrowing of vocabulary from indigenous languages, the development of locally characteristic collocations and ‘lexical bundles’ (Biber at al. 1999: 987ff.), but also grammatical nativization effects such as new word formation products or innovative verb complementation patterns (Schneider 2003: 249).

Phase 4 – Endonormative stabilization: This phase, which is characterized “by the gradual adoption and acceptance of an indigenous linguistic norm, supported by a new, locally rooted linguistic self-confidence” (Schneider 2003: 249), typically occurs after the country has gained political independence from Britain. The “newly achieved psychological independence and the acceptance of a new, indigenous identity result in the acceptance of local forms of English as a means of expression of that new identity” (p. 250). This new identity encompasses both STL and IDG strands, so that the differences between these groups become increasingly less important and tend to be downplayed. According to Trudgill, initially, such mixed, colonial varieties are typically extremely homogeneous (1986: 145), partly due to the koinéization that took place earlier in the STL strand (cf. Schneider 2003: 251), but, more importantly, because “in the phase of endonormative stabilization putting an emphasis on the unity and homogeneity of one’s own still relatively new and shaky identity is a natural socio-political move serving to strengthen internal group coherence” (p. 251).

Phase 5 – Differentiation: In this last stage, the New English variety is endonormatively stabilized. This marks a turning point and the new variety begins to diversify with the result that a wide range of regional and social dialects develop “as carriers of new group identities within the overall community” (Schneider 2003: 253). Differences between the STL and IDG strands tend to reappear as ethnic dialects, as de Klerk (1996) observes in the case of present-day South Africa.

Schneider emphasizes that these five stages describe the idealized developmental process which, however, need not apply in every single instance, as “boundaries and the chronology of stages may be fuzzy” (2003: 254). In the same paper he presents several case studies in which he determines for a number of varieties how far they have proceeded in the developmental cycle. For instance, he locates the variety of English spoken in Fiji in phase two, exonormative stabilization, and suggests that weak indications of further progress are already visible which are, however, not (yet) indicative of the structural nativization process setting in. Hong Kong is located in the third stage, nativization, and Singapore in phase four, endonormative stabilization. Malaysia and the Philippines prove
difficult to locate in either phase three or four. Schneider includes Australia and New Zealand as examples of varieties in stage five, diversification, although the sociopolitical and linguistic circumstances were different from those in the other varieties. The criticism put forward by Mesthrie and Bhatt is that “the incorporation of ‘dominion’ countries like Australia and New Zealand in a model that deals mostly with ESL countries like Fiji seems unwarranted” (2008: 35). Nevertheless, Schneider’s cyclic model is considered to be “the most elaborate account of the genesis of the New Englishes” (Sand 2005: 11). Following the publication of this ‘Dynamic model of the evolution of New Englishes’ linguists interested in World Englishes have recently begun to pay close attention to linguistic variation in New Englishes, focusing particularly on differences and similarities between these varieties, in order to understand the underlying processes of linguistic change they represent. Some issues discussed in such research are briefly mentioned in the next section.

2.1.3 Linguistic Variation in New Englishes and Underlying Principles

Linguistic variation is commonly attributed to one or more of the following factors, 1) language contact, 2) innovation, or 3) widespread trends and tendencies in varieties of English.

Variation through language contact means that a given variety of English may be directly influenced by Englishes which are historically or culturally dominant, e.g. British English due to its former colonial supremacy in many English-speaking regions of the world, or American English due to the status of the United States as a world power with extreme economic and cultural significance. Alternatively, a variety of English may be subject to substrate influence, i.e. it may display features deriving from the grammar or lexicon of a background language which functions as a widely used linguistic code in the given region. If a certain feature develops in a given variety of English but not in others with comparable sociolinguistic conditions and that feature cannot be attributed to language contact, then linguists speak of innovation. For instance, innovation is often assumed to be the most plausible reason for uncommon features emerging in isolated regions where dialect contact or language contact is very scarce or totally absent.

See Mesthrie and Bhatt’s (2008: 35f.) discussion of some further shortcomings of Schneider’s (2003) model and the stages proposed therein.
Features which develop independently in different post-colonial varieties of English around the world without language contact being a possible explanation are most likely caused by tendencies which are common to all vernacular varieties of English, i.e. so-called vernacular universals, or by angloversals. Chambers defines vernacular universals as “a small number of phonological and grammatical processes [which] recur in vernaculars wherever they are spoken” (2004: 128). According to Chambers the finding of these features in varieties around the world together with the observation that they occur not only in working-class and rural language but in all sorts of vernaculars, including child language, pidgins, creoles and interlanguage varieties (Chambers 2004: 128) rules out diffusion as a plausible explanation. Instead, “they appear to be natural outgrowths, so to speak, of the language faculty, that is, the species-specific bioprogram that allows (indeed, requires) normal human beings to become *homo loquens*” (Chambers 2004: 128).

However, the present study concentrates on (near-) standard varieties of English, as only those are covered in the components of ICE-Jamaica. In the centre of attention are instead angloversals, i.e. “joint tendencies observable in the course of the standardization of postcolonial varieties of English which cannot be explained historically or genetically” (Mair 2003: 84, translation taken from Kortmann and Szmrecsanyi 2004: 1150).

Szmrecsanyi and Kortmann (2008) draw attention to some terminological and typological issues and note that, for instance, the term ‘vernacular universal’ should be used with great care, as it bears the implication of denoting tendencies that occur not only in vernacular varieties of English but in vernaculars of all languages, so that research focusing on English alone cannot in fact make reliable assumptions about vernacular universals. The authors instead propose a more fine-grained typological classification of ‘versals,’ distinguishing between:

i. **GENUINE UNIVERSALS** (e.g. *all languages have vowels*);

ii. **TYPOVERSALS**, i.e. features that are common to languages of a specific typological type (e.g. *SOV languages tend to have postpositions*);

iii. **PHYLOVERSALS**, i.e. features that are shared by a family of genetically related languages (e.g. *languages belonging to the Indoeuropean language family distinguish masculine and feminine gender*);

iv. **AEROVERSALS**, i.e. features common in languages which are in geographical proximity (e.g. *languages belonging to the Balkan sprachbund have finite complement clauses*);

v. **VERNACULAR UNIVERSALS**, i.e. features that are common to spoken vernaculars (e.g. *spoken vernaculars tend to have double negation*);

---

20 These ‘vernacular roots’ are described in much detail in Chambers (2003: 266-270).
21 The data used and some characteristics of ICE-corpora are described in chapter 4.
22 Simo Bobda (e.g. 1998b) labels such tendencies applying to postcolonial Englishes ‘New Englishisms.’
vi Features that tend to recur in vernacular varieties of a specific language: ANGLOVERSALS, FRANCOVERSALS, etc. (e.g. in English vernaculars, adverbs tend to have the same morphological form as adjectives);

vii VARIOVERSALS, i.e. features recurrent in language varieties with a similar socio-history, historical depth, and mode of acquisition (e.g. L2 varieties of English tend to use resumptive pronouns in relative clauses) (Szmrecsanyi and Kortmann 2008: 2f.)

It seems that Simo Bobda’s (1998b) ‘New Englishisms’ and Mair’s (2003) ‘angloversals,’ denoting tendencies that apply specifically to postcolonial Englishes, correspond to ‘varioversals’ in Szmrecsanyi and Kortmann’s terminology, while the usage of the term ‘vernacular universal’ in the literature corresponds more closely to Szmrecsanyi and Kortmann’s ‘angloversals’. The present study investigates varieties of English only and is interested in tendencies of postcolonial L2 varieties as opposed to L1 varieties, i.e. varioversals. However, in keeping with much of the literature on the topic, the term ‘vernacular universal,’ when it is used throughout this book, is intended to mean vernacular varieties of English around the world, and the term ‘angloversals’ is used as intended by Mair (2003).

In recent years, scholarly interest in the search for potential explanations of features shared by many or all post-colonial Englishes and attempts to weigh the influence of substrate influence versus tendencies applying to all post-colonial varieties, i.e. angloversals (Mair 2003) has been huge indeed. This is reflected in the fact that entire conferences have been dedicated to these issues, e.g. a World Englishes conference on ‘Vernacular Universals versus Contact-Induced Change’ which took place in Mekrijärvi, Finland in 2006, as well as in the wealth of literature available on the topic, including a special issue of the Journal of Sociolinguistics entitled ‘Investigating Change and Variation Through Dialect Contact,’ a number of edited volumes, e.g. ‘Vernacular Universals and Language Contacts’ edited by Filppula, Klemola and Paulasto (2008), but also works such as Kortmann et al. (2004), Kortmann and Szmrecsanyi (2004), Szmrecsanyi and Kortmann (2008), Chambers (2000, 2001, 2003 and 2004), Sand (2005), Nevalainen (2006), and many more. In contrast to most of these scholars, who focus on substrate influence versus vernacular universals as the underlying processes of linguistic change in high-contact varieties, Trudgill (2006) addresses a question that he also finds worthy of academic endeavour: “[H]ow are we to explain, in sociolinguistic terms, developments that occur in low contact languages and dialects?” According to Chambers (2004: 134) Tristan da Cunha English (see Schreier 2002) offers itself as a suitable candidate for such studies of low contact varieties, as it is
one of the most isolated inhabited territories in the world and can even be considered “the sociolinguists’ Galapagos” (Chambers 2004: 134). On the background of the large amount of research and her own work, Thomason comes to the conclusion that “drawing a dichotomy between proposed ‘vernacular universals’ and contact-induced change is not a good idea, because many linguistic changes involve both various processes of contact-induced change and universal tendencies of various kinds” (Thomason 2006).

The present study intends to provide a clearer picture on how exactly such universal tendencies and language contact effects influence educated Jamaican English by first investigating variety-internal stylistic variation in the patterns of subject-verb agreement across a number of spoken and written registers in ICE-Jamaica (chapter 4), and then by comparing Jamaican English with two L1 varieties and two postcolonial L2 varieties with regard to existential constructions (chapters 5 and 6). The next section provides a brief summary of previous research on linguistic variation in Jamaica, particularly in the area of morphosyntax. The first part illustrates the whole range of morphosyntactic variation covered in the creole continuum and the second part gives a short summary of morphosyntactic characteristics of educated Jamaican English.

### 2.2 Linguistic Variation in Jamaica

#### 2.2.1 The Creole Continuum

In order to provide a theoretical framework accounting for the extremely high amount of linguistic variability in the Anglophone Caribbean, especially in Jamaica and Guyana, David DeCamp (1961: 82) proposed the concept of the creole continuum which he later developed further (1971; see also Rickford’s (1987) influential work on the creole continuum in Guyana) and which is nowadays accepted as a suitable model of synchronic variation in the Jamaican context. DeCamp’s model emphasizes the non-discrete transition from one end point, i.e. basilectal creole, through an indefinite number of intermediate varieties to the other end point of the spectrum of linguistic forms, the regional standard variety of the lexifier English, or acrolect. As Sand observes, this non-discrete nature of variation in the Anglophone Caribbean raises the important question “of whether the ends of the continuum constitute two different linguistic systems, or whether there is only one underlying system for the whole continuum” (1999: 57). This question was controversially discussed in the literature in the years to follow.
Although DeCamp himself admits that “[t]he linguistic variation in Jamaica is, of course, not literally a continuum” (1971: 354), he nevertheless justifies the validity of the continuum model:

> By calling it a continuum I mean that given two samples of Jamaican speech which differ substantially from one another, it is usually possible to find a third intermediate level in an additional sample. Thus it is not practicable to describe the system in terms of two or three or six or any other manageable number of discrete social dialects. (DeCamp 1971: 354)

Proponents of this view are, for instance, Craig (1980: 111f.), Romaine (1988: 158) and more recently Sand (1999: 177), Patrick (1999 and 2004: 410) and Mair (2003). However, there are also linguists (e.g. Holm 1988 and Winford 1988: 96ff.) who prefer to consider creole and English to be separate linguistic systems, where “the lines of demarcation are very hard to draw” (Bailey 1966: 1) due to “extensive cross-interference” (Bailey 1966: 1) and code-mixing of the two systems (cf. also Lawton 1980). The latter view mirrors widespread perceptions of the linguistic situation by the native population. As DeCamp notes, “[m]any Jamaicans persist in the myth that there are only two varieties: the patois and the standard” (1971: 350), as for instance in a situation of diglossia or bilingualism (cf. Deuber 2009a: 2). The intermediate varieties are seen as either belonging to the creole with extensive influence from English or vice versa. In fact, Bailey seems to have no doubt as to the existence of two separate systems but expresses “a compelling need for some system […] to determine whether a given specimen of language was standard with incursions from the creole, or creole with incursions from the standard” (Bailey 1971: 342).

However, Romaine (1988: 158f.) shows with Guyanese examples that it is practically impossible to determine the line between these varieties:

- \(mi\ gii\ am\)
- \(mi\ bin\ gii\ am\)
- \(mi\ bin\ gii\ ii\)
- \(mi\ di\ gii\ ii\)
- \(mi\ di\ gii\ hii\)

\(^{23}\) Patrick in principle agrees with the model of the continuum, emphasizing that “mesolectal grammar does not result from improvised mixing or code-switching between polar varieties” (2004: 410). However, he notes that “[d]espite the defining presence of English elements, which mark it off clearly from the basilect, the mesolect shares with the latter many constraints, structures and organizing principles which are not generally characteristic of native dialects of English” (Patrick 2004: 410) and hence comes to the conclusion that the mesolect is Jamaican Creole and cannot be genetically related to English (see Patrick 2004).
In the linear succession of examples from basilectal *mi gii am* to Standard English *I gave him* it is not possible to identify a discrete point where one variety stops and the other begins.\(^{24}\) This extreme difficulty is often put forward as one of the main arguments by opponents of the two-system view, which, as a result, is no longer widely defended in the field.

However, the continuum model, too, entails some complications. According to Sand, “one-system proponents […] sometimes run into difficulties when trying to account for all the necessary conversion rules needed to switch from one end of the continuum to the other” (1999: 58). Moreover, the speakers’ linguistic competence, i.e. the range of forms they cover, is hard to locate on the continuum, firstly because speakers understand more forms than they are capable of producing (cf. Winford 1988: 99), i.e. their passive competence is much larger than their active competence (cf. also Sand 1999: 57 footnote 62), and secondly because individual speakers cover different ranges of the continuum, i.e. they have different repertoires of linguistic variants on which they draw to modify their speech according to situational factors. Hence “one speaker’s attempt at the broad patois may be closer to the standard end of the spectrum than is another speaker’s attempt at the standard” (DeCamp 1971: 350).

In order to account for this non-discrete and linear variation, DeCamp introduced implicational scaling as method for investigating the creole continuum (cf. Sand 1999: 53) and illustrates this for six variables and seven speakers, all of whom participated in his

---

\(^{24}\) See Sand (1999: 57f.) for a more extended discussion of these examples and the problematic identification of variety boundaries.
survey of Jamaican speech (cf. DeCamp 1971: 355ff.). By means of this analytical tool he orders the speakers along a linguistic continuum according to the variants they used. The original mini-continuum presented by DeCamp (1971: 355ff.) was later modified for easier reading by several scholars, e.g. Rickford (1987). Rickford identified as fundamental the two characteristics (non-) discreteness, i.e. the notion that “variation across speakers is extremely fine-grained,” and unidimensionality, i.e. the notion that the linguistic variation “can be ordered in terms of a single dimension” (Rickford 1987: 22), usually identified as “standardness versus creoleness” (cf. Patrick 1999: 11).  

<table>
<thead>
<tr>
<th>Creole ‘C’:</th>
<th>/d/</th>
<th>/t/</th>
<th>pikni</th>
<th>no ben</th>
<th>nana</th>
<th>nyam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaker 4</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Speaker 3</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>e</td>
</tr>
<tr>
<td>Speaker 7</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>Speaker 2</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>e</td>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>Speaker 6</td>
<td>C</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>Speaker 1</td>
<td>C</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>Speaker 5</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>English ‘e’:</th>
<th>/ð/</th>
<th>/θ/</th>
<th>child</th>
<th>didn’t</th>
<th>granny</th>
<th>eat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaker 4</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Speaker 3</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>e</td>
</tr>
<tr>
<td>Speaker 7</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>e</td>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>Speaker 2</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>e</td>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>Speaker 6</td>
<td>C</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>Speaker 1</td>
<td>C</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>Speaker 5</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
</tr>
</tbody>
</table>

Source: Patrick (1999: 7, adapted from DeCamp 1971)

**Table 1: Implicational scale – variable features of Jamaican Creole and Jamaican English**

The table illustrates the variation in the speech of seven individual speakers between the polar end points of the continuum using a mixed set of invariant phonological, lexical and syntactic features. Of the phonological features, /ð/ and /θ/ represent Standard English pronunciation, and /d/ and /t/ are the corresponding creole variants. Similarly, the past marker no ben corresponds to English didn’t and the creole lexemes pikni, nana and nyam correspond to English child, granny and eat, respectively. What the table shows is that the two-system view postulating separate dialects is not capable of adequately portraying linguistic reality. In fact, the table indicates that speakers can be ordered on a continuum according to the linguistic forms they use. While speaker 4 produces creole forms in each case and uses none of the corresponding English alternatives, the reverse is true for speaker 5. In addition, the table provides evidence that variation is systematic and thus predictable. The presence of a particular feature in a speaker’s usage can predict the occurrence or

---

25 It is Patrick’s slightly adapted version of DeCamp’s (1971) continuum which is reproduced as Table 1. The following summary only touches upon some of the issues discussed in more detail in Patrick (1999: 6ff.). He gives a more thorough account on the creole continuum and the two fundamental characteristics of variation in Jamaica, (non-)discreteness and unidimensionality.
absence of other features. For example, speakers who use the lexeme pikni rather than child can be expected to favor the creole phonological forms /d/ and /t/ over their English counterparts, and speakers who use the lexeme child can be expected to produce the English forms didn’t, granny and eat rather than the creole alternatives. Although DeCamp has been criticized for under-representing the range of the continuum controlled by individual speakers in this model, i.e. the range of variants used by individual speakers for stylistic purposes, the table nevertheless proves an important point, namely that the gradual transitions within the continuum are not random but systematic.

While the table accounts for inter-speaker variation across the whole (social) continuum, it is important to note here that the present study is based on corpora belonging to the International Corpus of English project, in which only adult speakers (age 18 and older) are included who have at least secondary education or, alternatively, whose public status (for instance as politicians, broadcasters, etc.) justifies their inclusion. This is because in contrast to Jamaican Creole, which is the “national vernacular shared across social boundaries” (Patrick 1999: 272), competence in English is tied to the speaker’s position on the (social) continuum (cf. Deuber 2009a: 3), and English is the variety that the ICE project aims at documenting. Hence only the upper mesolectal and acrolectal range of the Jamaican creole continuum can be taken into account in the present study of variety-internal variation in Jamaican English. What follows in the next section is a summary of some (but by no means all) morphosyntactic peculiarities of educated Jamaican English identified in previous research.

2.2.2 Morphosyntactic Characteristics of Jamaican English

In her study ‘The English we speaking’ - morphological and syntactic variation in educated Jamaican speech, Deuber (2009a) thoroughly investigates stylistic variation in Jamaican English on the basis of 40 conversational texts from the spoken section of ICE-Jamaica. The guiding question Deuber addresses is

26 Deuber (2009) focuses on the question of stylistic variation in her analysis of 40 conversational texts in ICE-Jamaica.
27 The International Corpus of English and some of the principles of compilation, including the choice of speakers, are described in section 4.2.1.
28 See Sand (1999) for a detailed description of lexi-co-semantic (pp. 77-110) and morpho-syntactic (pp. 111-150) characteristics of Jamaican English. Christie (2003: 14-20) briefly summarizes some phonological and lexical differences between Jamaican and British English and takes into consideration areas of influence of American English. Readers interested in sociophonetic variation in Jamaican English are referred to the recent and thorough investigation by Rosenfelder (2008).
whether the creole continuum model can account for the way speakers like these [i.e. highly educated Jamaicans, S.J.], who have a full command of acrolectal Jamaican English and tend to be proficient in Jamaican Creole as well, make use of the range of varieties available to them. (Deuber 2009a: 1)

Deuber carries out a quantitative analysis of a number of morphological and syntactic features and then compares the results obtained in her analyses to those presented by Sand (1999) in her study on more formal radio and newspaper usage in Jamaica, in order to broaden the perspective to a wider range of speech situations. In a qualitative analysis of inter- and intratextual variation Deuber discusses the various ways in which Jamaicans adapt their usage to situational contexts. Deuber comes to the conclusion that, indeed, the spoken language practices of these speakers can be more appropriately described in terms of a continuum from more English to more Creole ways of speaking than in terms of the diglossia model which Mair (2002) has found to adequately describe most written practice. (Deuber 2009a: 47f.).

Some of the more frequent morphological and syntactic features attested in the corpus of educated Jamaican English are summarized below. However, the analysis of stylistic variation and Jamaicans’ use of creole features for rhetorical purposes that Deuber describes in much detail in her paper will be left aside here.

One feature that Deuber analyses quantitatively is the syntax of direct wh-questions, a relatively high number of which are found not to correspond to standard usage. Her analysis shows that in cases where the verb phrase in the equivalent declarative sentence does not contain a form of be (functioning as auxiliary or main verb) the corresponding questions frequently occur without ‘do-support’ as in (1).

(1) Which school you went to (S1A-029)

There are also instances in the data of questions in which subject and operator are not inverted, e.g. (2), and questions with zero copula (3).

(2) So what you’re doing for summer (S1A-032)
(3) Where in Clarendon you going (S1A-036)

In contrast to Sand (1999), who finds such creole-type questions to be rare and stylistically marked in her media texts, Deuber finds them to be one of the most frequent non-standard

---

29 The summary is based on Deuber’s (2009) study unless indicated otherwise. Features which occur but are very rare in her data are neglected here, and the reader is encouraged to consult Deuber’s paper for more details.
features in her conversational data and thus considers them to be a feature of informal educated Jamaican speech.

With regard to zero copula, a feature which Rickford (1999: 146) found to be frequent in DeCamp’s (1960) four Jamaican Creole stories, Deuber (2009a) reports few cases in pre-locative (5%) and pre-adjective contexts (4%) and somewhat higher proportions in present and past active progressives (9%) and going-to-futures (14%). She describes the latter two contexts as the least marked ones for zero copula. Deletion of copula in these contexts is, however, not restricted to Jamaican English but also occurs in British and American English in some cases, especially before gonna (cf. Biber et al. 1999: 1107 and Deuber 2009a: 15).

Another characteristic of Jamaican English is variable past marking, i.e. the alternation between inflected verbs and the zero form. While unmarked past participles are hardly attested at all in Deuber’s 40 conversational texts, non-participial verbs with past reference are more frequently uninflected. Deuber notes a marked difference between her own and Patrick’s mesolectal data30 from the Kingston neighbourhood Veeton not only with regard to the overall proportions of marked and unmarked past reference verbs but also in the inflection rates of various verb categories (cf. Patrick 1999: 231). In the comparison of her own results with those obtained by Patrick (1999) in mesolectal and Sand (1999) in more formal text types she finds evidence of an “acrolectic range in which unmarking of past reference verbs becomes more common as one goes from formal to less formal situations” (Deuber 2009a: 19).

Moreover, some instances were found of verbs with a preverbal past marker, as in (4), but these were much rarer than in Patrick’s mesolectal data.

(4) I feel kind of good to you know finish my degree and I finished in a few years since I did come here (S1A-015)

Although Deuber’s study also deals with agreement marking on verbs, finding it relatively common for verbs not to agree with their subjects, particularly in the third person singular, and providing evidence that Jamaican usage cannot be satisfactorily accounted for by general tendencies in informal spoken English, this phenomenon will be left aside here as her results are discussed in a later section.

---

30 Despite methodological complications arising from the lack of correspondence between the tense and aspect categories in creole versus English, the referential ambiguity of zero forms of verbs, and the resulting necessity for Deuber to modify her analysing procedure, she ensures the comparability of her and Patrick’s results by analysing the same verb categories as are distinguished by Patrick (1999).
Main verb negation is the next variable discussed by Deuber. She reports some instances of verbs which are negated by the basilectal preverbal negator no (cf. Patrick 1999: 199), as in (5), but notes that this is a rare form in her data.

(5) Me no think so (S1A-021)

Much more frequent are occurrences in the corpus of the contracted negative form of do with third person singular subjects (44%), e.g. he don’t, which Deuber attributes to influence of mesolectic Jamaican Creole, where the invariant negator duon is commonly used (cf. section 4.3.1). As Deuber shows, speakers also occasionally rely on the creole form neva as a marker of negative polarity in Jamaican English. There are two types of meanings of never in past reference. According to Deuber,

- it can indicate that something did not happen or obtain (a) on any occasion in a period of past time (which may lead up to the present or be completed) […];
- at a particular point of time or within a period of time in the past (leading up to the present or completed), but without the element of ‘on any occasion’ […]. (Deuber 2009a: 26)

Sentence (6) represents an example of the first type, and an illustration of the second meaning is given as (7).

(6) You never heard it in like songs (S1A-031)
(7) When him did born him never have no teeth (S1A-011)

While the first type occurs in the creole and in Standard English, the second type is rather common in Jamaican Creole but it is also widespread in other non-standard varieties of English (Kortmann and Szmrecsanyi 2004: 1154). According to Patrick (1999: 202), the creole form neva is generally followed by unmarked verbs and, indeed, Deuber (2009a: 26) finds that the occurrence of never reduces the probability of past marking in the following verb. As never + noun occurs especially in clusters of creole features, Deuber concludes that the creole pattern exerts an influence on English with regard to this feature.

As concerns noun morphology, Deuber reports a rather low number of nouns lacking plural inflection (about 2%) and attributes this to the fact that the plural –s marking is common also in the mesolect. A feature that Deuber finds more pervasive is unmarked genitives. Of all genitives in her data, 11.8% are unmarked (6 out of 51 tokens), e.g. the person friendship (S1A-033), compared to just one instance reported by Sand (1999: 139).
Hence Deuber concludes that “lack of genetive –s is indeed considered more of a Creole feature in the Jamaican context than unmarking of plural nouns” (Deuber 2009a: 27).

The last variable investigated is pronoun morphology and Deuber’s results indicate that personal and possessive pronoun use is generally based more on the Standard English rather than the creole forms, although some creole forms do occur in her corpus. She finds that the same creole pronoun is occasionally used for subject, object and possessive functions, e.g. first person singular me, third person singular masculine him31, and third person plural them, and postulates a cline among the creole forms “from those which are more salient to those which differ least from English” (Deuber 2009a: 28). She further notes that “the JamC [i.e. Jamaican Creole, S.J.] forms used most tend to be those whose phonetic shape differs least from their English counterparts” (Deuber 2009a: 28).

Some of the above-mentioned ‘typical’ features attested in Deuber’s data are not exclusive to Jamaican English but occur also in other non-standard varieties of Englishes (Kortmann and Szmrecsanyi 2004: 1154) and especially New Englishes (such as the second language varieties included in ICE), e.g. the lack of inversion or auxiliaries in wh-questions. However, as Deuber notes in her conclusion, the peculiar pattern of features revealed by her data and the way these are used can only be accounted for by influence from the creole substrate (2009a: 47).

As the focus of attention now turns to one particular morphosyntactic variable, namely subject-verb agreement, the complex web of influences on educated Jamaican English will be explored from different angles. The next chapter will give an overview of previous research on subject verb agreement, and this will be followed by a thorough analysis of both intra-variety variation and inter-variety variation.

31 The creole pronoun is used most frequently in the third person singular masculine context.
3 Previous Research on Subject-Verb Agreement

Agreement is a widespread phenomenon in the world’s languages. It is pervasive in some languages and largely absent in others. English is an analytic language in which grammatical information is rarely indicated by means of inflection. One of the few morphosyntactic features where this is the case is subject-verb agreement, which is the focus of the present study. As there is no consistent terminology in the literature, I will begin with a brief clarification of the terms that will be used throughout this book.

Up to the present, there has been no clear separation of the terms 'agreement' and 'concord' in the literature. They are often used synonymously, for instance by Hundt (2006) and Quirk et al. (1985), or in rather contradictory ways. While Bloomfield (1933: 191-194) interprets 'agreement' as the superordinate term which includes 'concord' as a subcategory, Greenberg (1978: 50) applies them in the opposite way, using 'concord' as the wider term. Corbett (2006: 6) even points to the fact that some linguists use 'concord' exclusively for agreement within the noun phrase, while they label all other cases 'agreement.' In front of such a background of inconsistent application of the terminology, ‘agreement’ will be the term used consistently throughout this book for all types of such phenomena. Similarly, cases of grammatical ‘non-concord’ or ‘discord’ will be called ‘non-agreement.’

Agreement is defined as “the relationship between two grammatical units such that one of them displays a particular feature (e.g. plurality) that accords with a displayed (or semantically implicit) feature in the other” (Quirk et al. 1985: 755). However, it is not sufficient if the two grammatical units happen to share properties. There has to be covariance, that is, “the sharing must be systematic” (Corbett 2006: 4). The basic rule of agreement is expressed by Quirk et al. (1985: 755) in very simple terms: “A singular subject requires a singular verb […]. A plural subject requires a plural verb.” To give an example of agreement, in (1) both subject and verb are plural, while in (2) the subject is singular, but the verb is not marked for singular, and in (3) the subject is plural but the verb has singular form. Examples (2) and (3) therefore represent cases of non-agreement.

(1) Uhm they teach you the science behind electronics (S1A-046)

(2) A student represent all the the committees on the boards (S1A-059)

32 The code in brackets refers to the corpus text in ICE-Jamaica from which the example was taken; in this case the source was text 046 of the text category S1A, i.e. ‘direct conversations’ from the spoken section. See chapter 4.2 for a more detailed description of the data used in the present study.
(3) ...even as additional things has been found to ease the plight of the unemployed… (S2B-017)

Normally, the subject noun phrase is the element which triggers, or ‘controls’ agreement and thus determines the form of the target, i.e. the verb. In (4) below, the NP those comments functions as the controller of agreement, while the verb come is the so-called ‘target’ which agrees in number with the controller.

(4) Those comments come largely from two sets of people (S1A-048)

There are, however, cases in which not the item in subject position but another element controls agreement. For instance, in existential there + be constructions the subject position is taken by the dummy there, and it is instead the complement following the verb which functions as the controller of agreement. Example (5) is considered to be an instance of agreement, as the contracted verb form agrees in number with the complement a company, while in (6) the complement four books is clearly plural, but the past tense verb form was is singular and thus does not agree grammatically33 with the controller.

(5) Uhm when we look at uhm our neighbours to the north in Saint Ann there’s a company by name of Walkerswood (S1A-049)

(6) I think there was four books (S1A-064)

Subject noun phrases consisting of a quantifier, such as a number of (7) or a lot of (8) are another case in point, since it is not the syntactic head of the respective prepositional phrase which controls agreement, i.e. number or lot, but the noun whose number is thus specified. Examples (a) are instances of agreement and examples (b) are cases of non-agreement.

(7a) She explained that each academic year a number of students enter the first form unprepared for high school work (W2C-004)

(7b) When attempting any kind of movement as regards to passives a number of features or properties becomes evident (W1A-017)

(8a) I know a lot of people say the university does not give… (S1A-057)

(8b) A lot of the kids in my class (not classes) doesn’t like me (W1B-003)

33 Potential notional interpretations of numeral + noun combinations, e.g. ‘four books’, as denoting a single entity will be considered in a later section, the point of departure here is grammatical number.
Whereas phrases involving *a number of* are generally followed by plural count nouns (7a and b), the controller of phrases involving *a lot of* can be either a count noun (8a and b) or a non-count noun (8c) and the phrase can thus be plural or singular.

(8c) Mr Taylor says **there is a lot of potential** for the local industry to improve both its levels of efficiency and production (S2B-005)

The domain of agreement, i.e. the syntactic environment in which agreement occurs, is normally the clause, but in some cases it can be the whole sentence (9) or even extend beyond the sentence in case of pronoun-antecedent agreement (10), the latter of which will, however, only play a minor role in the present study.

(9) I’ve spoken to **a lot of people** who **seem** to think that uhm these aren’t negative moves… (S1A-048)

(10) I have **a little niece**. **She’s** eight (S1A-052)

The standard view on subject-verb agreement in English is that verbs (except modal auxiliaries) normally agree in person and number with their respective subjects. However, as Hudson (1999) demonstrates, the inflectional system in English is reduced, in that verbs only display agreement features in present tense and agree with their subject merely in number but not in person. Third person singular subjects are the only controllers which lead to inflectional –s marking in the verb. The verb *be* is exceptional, as it agrees with its subject in number and person, i.e. 1st person *I am* versus *we are*, 2nd person *you are* for singular and plural, and 3rd person *he/she/it is* versus *they are* in present tense, but *be* also displays agreement features in past tense by means of the *was – were* distinction.

The subject-verb agreement system is, however, more complex than Quirk’s general rule above suggests, as there are a great number of exceptional subjects whose grammatical number is difficult to specify. Besides count nouns which are overtly marked for plural by adding an –s to the noun, e.g. *a book* versus *three books*, other nouns are marked for number by means of vowel change, for instance *foot* versus *feet* or *woman* versus *women* or a change in the diphthong, as in *mouse* versus *mice*. There are also nouns whose singular and plural forms are identical and not inflectionally marked, as in *sheep* or *fish*. Some nouns are marked for plural by means of endings that survived from Old English, for example, *child* versus *children* or *ox* versus *oxen*. Moreover, some non-count nouns have no plural form, for instance *advice, knowledge or information*, while others are inherently plural and do not have a corresponding singular form, such as *trousers* or *scissors* (cf.
The noun *people*, meaning ‘several individuals,’ normally takes plural agreement but when it is used in its singular form it has a different, more specific, meaning, denoting a whole group of persons belonging to the same community.\(^{34}\) Thus, with a change in number there is also a semantic change. Rastall (2007) shows that adjectives, when used as nouns, are plural in meaning without being marked for plural, as in *The poor are always with us* or *The good die young.*\(^{35}\) Nouns like *economics* or *politics* seem to be plural due to their ending in –s, but they can be interpreted as singular or plural in different contexts (Rastall 2007: 51).

Moreover, English also has a large number of so-called collective nouns, for instance *family*, *government* or *administration*, which have singular form but can be interpreted as singular or plural depending on whether the focus is on the whole group or its individual members. While these nouns lack a morphological plural marker, i.e. are singular in form, and therefore require a singular verb for grammatical agreement, they are sometimes followed by plural verbs due to the notion of plural number, in which case grammarians speak of ‘notional agreement’ (cf. Quirk et al. 1985: 757). Sand (2005: 156) reports findings from previous research on agreement with collective nouns in native Englishes, such as American, British English or New Zealand English (e.g. Hundt 1998: 80-89, Depraetere 2003, Levin 1998, 1999, 2001 or Bauer 1994)\(^{36}\), which all found that agreement with collective nouns is variable in the respective varieties, but that “the general trend appears to be singular concord.” In other words, grammatical agreement seems to be the preferred form. The conclusion all these linguists arrive at is that “singular concord is used unless there are clear reasons for using the plural” (Biber et al. 1999: 181), that is, plural “denotes more than one” (Quirk et al. 1985: 297), but singular is not restricted to “that which ‘denotes one’” (Quirk et al. 1985: 297, quoted in Rastall 2007: 51).

However, occasionally it is neither the morphological marking of the subject nor the semantic notion which determines the number of the verb, but the distance between the subject and the verb. This phenomenon is called ‘attraction’ (see Quirk et al. 1985: 757) or

---

\(^{34}\) Rastall shows that even in the singular, the noun *people* contains a sense of plurality (Rastall 2007: 51 and footnote 2).

\(^{35}\) Some of the examples presented in this section were taken from Rastall (2007: 51-55). Quirk et al. (1985: 755-767) and Biber et al. (1999: 180-190) give a detailed description of many more such complications in the agreement system, such as coordinated subjects.

‘proximity concord’\textsuperscript{37} (Biber et al. 1999: 189f. and Francis 1986: 311) and includes cases in which the verb does not agree with the head but with a nearer noun in a complex noun phrase. Mair (2002: 52) presents the following example (11) from ICE-Jamaica:

(11) The arrangements for each level I think is reasonable

Here the verb \textit{is} agrees not with the plural head noun \textit{arrangements}, but instead with the local noun \textit{level}. A main finding of experimental psycholinguistic research on “attraction errors” (Hartsuiker, Antón-Méndez and van Zee 2001: 547) is that agreement errors occur more frequently when the local noun is plural than when it is singular. This is in line with Francis’ observation that the most frequent pattern is “Nsg + Prep + Npl + Vpl” (1986: 314), i.e. proximity agreement more often causes a change from singular to plural, as in (12), than vice versa.

(12) The heads do not give reference, it is the \textit{combination} of the head and the other items that give reference… (W1A-011)

This finding can be plausibly explained by the trend toward morphological simplification, i.e. the tendency of leaving out inflection rather than adding it, but in this example from Jamaican English the instance of non-agreement may be an effect of language contact, as the lack of s-marking on third person singular verbs (and on plural nouns) is a pervasive feature of Jamaican Creole.\textsuperscript{38}

An important point to note with regard to proximity agreement is that proximity should be understood in syntactic terms, not simply in terms of the presence or absence of intervening constituents (see Hartsuiker, Antón-Méndez and van Zee 2001: 547). Bock and Cutting (1992) present two sentences (13a and b) and show that there is a difference in syntactic proximity between the subject and the respective local noun and verb.

(13a) The editor of the history books was/were...
(13b) The editor who rejected the books was/were…

While the subject of the first sentence is post-modified by a prepositional phrase and subject and local noun are located in the same clause, in the second sentence it is a clause which post-modifies the subject, so that subject and local noun are syntactically further apart.

\textsuperscript{37} As the author favours the term ‘agreement’ over ‘concord’ (see above for details), this phenomenon will be called ‘proximity agreement’ throughout this book.

\textsuperscript{38} Substrate influence in Jamaican English will be discussed in detail in later sections.
Psychologists (See Hartsuiker, Antón-Méndez and van Zee, 2001 for a summary of such work) found in a series of experiments that error rates were significantly higher when the local noun occurred in a post-modifying phrase, than when it was located in a clause. It is not the clause boundary but the number of phrasal nodes between the local noun and the highest node of the subject noun phrase that determines the proximity or distance.

Linguists greatly differ in their findings as to where proximity agreement is most likely to occur. Quirk et al. (1985: 757) report that “[p]roximity concord occurs mainly in unplanned discourse” and that “[i]n writing it will be corrected to grammatical concord if it is noticed” (see also Sand 2005: 156f.). Biber et al. provide support for this assumption as follows:

Such clear deviations from grammatical concord are mostly found in speech, where they are explicable from the psycholinguistic constraints of a limited short-term memory and the pressure of online construction of linguistic output. They are rare in writing though we do find proximity effects in special cases. (Biber et al. 1999: 189f.)

Francis (1986), however, comes to an altogether different conclusion in his investigation of proximity agreement, which is based on a casual collection of examples rather than quantitative corpus analyses. He finds that “proximity concord is occurring in the upper stylistic registers” (1986: 316), “mostly from edited English” (Francis 1986: 316). His results clearly contradict those presented by Quirk et al. (1985) and Biber et al. (1999), but evidence can easily be obtained which supports Francis’ view. In his investigation of written Jamaican English, Mair (1992) reports for newspaper usage that “uncorrected instances of the proximity principle in action are the most frequent grammatical peculiarity observed” and finds this assumption confirmed in student essays (2002). Jantos (forthc.) and Deuber (2009a), who analyse spoken Jamaican English, both find proximity agreement to be extremely rare in informal direct conversations. The study presented in chapter 4 intends to provide a clearer picture of proximity agreement by systematically investigating potential reciprocal effects of the mode of production and stylistic factors, especially the level of formality, on the frequency of proximity agreement in Jamaican English.

After reviewing the literature, Francis (1986: 311) summarizes the three main types of subject-verb agreement in English as follows:

Grammatical concord is morphologically based.
Notional concord is semantically based “according to the notion of number rather than… the actual presence of the grammatical marker” (Quirk et al. 1985: 757).
Proximity concord or “attraction” is positionally based.
While grammatical agreement is widely seen as the standard pattern, notional and proximity agreement is non-standard in the sense that it is not the default form. Such non-standard subject-verb agreement, while not being a categorical choice or the preferred form (Sand 2005: 198\textsuperscript{39}), has, however, been characterised as a widespread phenomenon that occurs in native as well as non-native varieties of English (see also Szmrecsanyi and Kortmann 2009). The underlying factors determining whether subject and verb agree in number may be different for different variety types, ranging from general tendencies toward simplification and regularization found in native varieties (cf. Trudgill 1990 and Hudson 1999 for dialects of British English), over partial loss of inflectional morphology due to language contact (cf. Schneider 2003) and the levelling of irregularities which has been attributed to “L2-acquisition difficulty” (Szmrecsanyi and Kortmann 2009), to the influence from substrate languages (Platt, Weber and Ho. 1984) or the focus on function rather than form in English as a lingua franca (ELF) contexts (Breiteneder 2007). Moreover, different varieties show preferences for either singular or plural verbs after collective nouns, thus bringing regional differences into the discussion. While not all these studies can be presented here in detail, the next section will be devoted to a summary of some main findings on subject-verb agreement in varieties of English around the world.

3.1 Agreement with Lexical Verbs

Research into subject-verb agreement with lexical verbs has mainly concentrated on native varieties of English, such as Irish English (Harris 1993), British English (Trudgill 1990), American English (Crawford 2005), or other European languages, for example Dutch (Hartsuiker, Antón-Méndez and van Zee 2001), Italian (Vigliocco, Butterworth and Semenza 1995) and others. Linguists have only recently begun to investigate subject-verb agreement in non-native or ‘outer-circle’ varieties of English, for instance Sand (2005) and Hundt (2006).

Research has shown that there is a lot of variation in the agreement patterns of different varieties of English. According to Trudgill (1990, quoted in Hudson 1999: 203 f., cf. also Trudgill 2002), there is a diachronic shift towards loss of subject-verb agreement that is visible in some British dialects. He shows that such simplification processes affecting

\textsuperscript{39} Sand’s conclusions are based on an investigation of non-standard agreement in Irish English, Indian English, Jamaican English, Kenyan English and Singapore English.
present-tense verbs are present in East Anglian English and dialects of the north and west of England. While in East Anglian dialects verbs in the present tense tend to have no -s-form and examples such as “He like her,” “She want some” or “That rain a lot there” occur quite frequently (Trudgill 1990: 94), the -s suffix can appear in all present-tense forms in many dialects of the west and north of England, for instance “I wants it,” “We likes it,” “They sees them” (Trudgill 1990: 94). Hudson also shows that subject-verb agreement is partially lost in many urban British dialects, for example in negative auxiliary verbs and the non-modals ain’t (14), used for haven’t/hasn’t and for aren’t/isn’t, and don’t (15) (1999: 203)

(14) He ain't working/worked at all.
(15) He don't work at all.

Apart from regional factors the type of subject has been found to determine agreement, for instance in Irish English, where the Northern Subject Rule applies (cf. Harris 1993). This means that agreement with personal pronoun subjects is different from agreement with common nouns, as shown in (16).

(16) Them eggs is cracked, so they are. (Harris, 1993: 155)

The third person plural pronoun they is given obligatory plural number. This means that the –s marked verb is ungrammatical in a sentence with a plural subject pronoun, while for other subjects agreement is variable (Hudson 1999: 204).

With regard to simplification of inflection in non-native Englishes, Platt, Weber and Ho (1984: 67) report that in New Englishes, i.e. contact varieties of English from around the world, for instance India, Jamaica, Sri Lanka, Singapore, and others, verbs are often not marked for 3rd person singular in the present tense, and that plural nouns frequently lack plural inflection (1984: 47). However, they do not go into much detail about possible reasons for such reduced inflection.

A large proportion of the research on subject-verb agreement has been published by psychologists and psycholinguists. Many of these scholars carried out experiments on the production of agreement errors in verbs following complex subject NPs, i.e. attraction effects which are here called proximity agreement (see above). In such experiments, not naturally occurring conversations are analysed but agreement errors are usually artificially
elicited. Research interests in studies of this kind have mainly concentrated on the following two issues:

1) The influence of syntax versus semantics, i.e. grammatical versus notional agreement, and

2) The question how the grammatical number of head and local noun respectively affect agreement, that is, how matching versus mismatching number features on head versus local noun affect agreement.40

The statement by Haskell and McDonald that “English speakers must produce subject-verb agreement for virtually every sentence they speak” (2003: 760) is symptomatic for much of the research carried out on subject-verb agreement, as it clearly shows the focus on spoken production. Spoken language has been analysed in much detail, while only few studies so far have thoroughly investigated agreement patterns in written language use. One of the few investigations that analyse both spoken and written language is presented by Hundt (2006), who explores subject-verb agreement with collective nouns in a number of varieties that can be classified as inner- and outer-circle Englishes according to Kachru’s concentric circles model of global English (cf. section 2.1.2 and Kachru 1997). Hundt uses the International Corpus of English as the basis from which she draws her data and uses British and New Zealand English as representatives of inner-circle varieties, that is native Englishes, and Philippine English (American-based) and Singapore English (British-based) as outer-circle varieties, that is English as a Second Language (ESL) varieties. The question Hundt intends to answer is whether agreement patterns in PhilE and SingE exhibit a close correspondence with their historically related inner-circle varieties[...] or whether the patterns of variation will provide evidence of divergence—a finding that might suggest either exonormative pressure from other inner-circle varieties[...] or the development of an endonormative native model. (Hundt 2006: 208)

Her results show that there is “not much evidence of a strong endonormative model so far” (Hundt 2006: 222) but that the exonormative influence is quite strong, and more so in PhilE than in SingE. Hundt finds influence from the substrate at the lexicogrammatical level in her data, but substrate influence does not play a significant role in agreement with collective nouns in acrolectal English of the outer-circle varieties. One of her main findings is that there is a “greater internal stylistic homogeneity” (Hundt 2006: 223) in the

40 The asymmetrical case of a singular head noun co-occurring with a plural local noun, i.e. a mismatch of number features, leads to most agreement errors.
outer-circle varieties than in the native Englishes. In other words, spoken PhilE and SingE appear to resemble their written norms more closely than spoken AmE or BritE. Hundt takes this as “an indication of a persistent exonor mative, inner-circle model” (2006: 223) and argues that “[i]f people speak more or less as they write, the abstract, underlying model toward which they aim is that of written Standard English” (2006: 223). She attributes the strong orientation towards this written standard to the “important role of the educational sector with (1) its emphasis of written production, (2) overt notions of “correctness,” and (3) a largely imported (written) norm as its yardstick for comparison” (Hundt 2006: 223).

As mentioned above, studies on Jamaican English found that theacrolectal variety is moving away from the historical norm. Sand, however, points to the fact that the "features deviating from international Standard English usage […] are not the preferred forms" (1999: 148) in the public formal domain of the press. Mair (2002: 36) reports that there is a high degree of variation in spoken language and postulates that “Patois is clearly the dominant shaping influence on spoken English in Jamaica,” not only as regards phonological features, but also in terms of syntax and lexis. He does not, however, go into detail about “the complicated interaction of the creole and standard English norms on the level of textual macro-structure” (Mair 2002: 37). In fact, there have been few comparative studies dealing with variation across registers in Jamaican English. Sand (1999), for instance, focuses on the mode of production in her analysis of radio and press texts, while Mair analyses one particular written register at a time, namely newspaper usage (1992) and student essays (2002). Systematic investigations of Jamaican data in terms of register characteristics or effects of situational factors such as formality are still rare and no study to date has carried out a comparison of spoken and written language with regard to these factors. The present paper intends to fill this research gap by systematically analysing three spoken registers differing in their relative levels of formality and by comparing these to corresponding written registers.

In her analysis of 40 direct conversations, Deuber (2009a) found that non-agreement of regular verbs with their respective subjects is relatively frequent (the rate being around 20 percent). She did not find evidence of the proximity principle or the semantic notion of the noun being strong explanatory factors in her data, rather, “simple unmarking of third person singular verbs is clearly the predominant phenomenon” (Deuber 2009a: 22), due to influence from the creole, where third person singular –s marking on verbs does not exist
While Trudgill provides an alternative explanation and attributes this lack of –s marking in Jamaican English and other varieties, e.g. Pitcairnese, to imperfect learning of a linguistic code by speakers involved in language contact generally, suggesting that “the loss of an irregular and typologically unusual form of personal marking is a typical example of the kind of simplification which takes place in contact situations involving adults” (2002: 104), Sand rejects this assumption on the grounds that the occurrence of these features is one stylistic alternative in the linguistic repertoire of speakers with a high proficiency in English and that these speakers “choose to alternate between morphologically marked and unmarked forms in certain contexts” (2005: 198). Sand (1999) found evidence for creole influence also in the use of quantifiers. An example is *much*, which in Jamaican English often occurs with nouns that function as count nouns in International Standard English. Sand attributes this to the fact that *much* can be used irrespective of count versus mass-noun distinctions in Jamaican Creole (cf. Bailey, 1966: 30). However, Sand (1999) suggests that reclassifications of count and mass nouns are not restricted to the quantifier, but that there is a general, though slight, tendency in Jamaican English to extend notional concord. For lack of evidence Sand cannot determine whether or not such reclassifications are a stable feature of Jamaican English. Deuber’s results support Sand’s assumption in that, in Deuber’s data, there are frequent examples of *everybody* or phrases with *every*, which can be interpreted as notionally plural, being followed by a plural verb (Deuber 2009a: 22).

Proximity agreement seems to be an extremely infrequent phenomenon in private conversations (cf. Deuber 2009a and Jantos forthcoming), but was found rather frequently in Sand’s (1999) spoken and written media texts. Sand did not find any influence of the mode of production on proximity agreement, as the percentage was roughly the same in Jampress (written) and Jamradio (spoken), with 10.1 and 10.8 percent respectively. This result is in stark contrast with Quirk’s observation that “Proximity concord occurs mainly in unplanned discourse. In writing it will be corrected to grammatical concord if noticed” (Quirk et al. 1985: 757). Deuber’s and Sand’s results lead to the conclusion that the level of formality might exert a strong influence on this type of agreement. Mair relativizes this hypothesis, since, in his investigation of newspapers, “uncorrected instances of the proximity principle in action are the most frequent grammatical peculiarity observed” (1992: 82), but his analyses of student essays, which have a lower average degree of formality, yield similar amounts of proximity agreement (2002: 52). The question whether
it is indeed the level of formality which influences proximity agreement or whether other factors, such as the spoken versus written dichotomy, might be more crucial will be addressed in chapter 4.

3.2 Agreement with the Verb BE

With regard to the verb *be* studies so far have focussed on native varieties of English, such as Australian English (Korhonen 2007), New Zealand English (Hay and Schreier 2004), American English (Crawford 2005), British English (Tagliamonte 1998) and Canadian English (Meechan and Foley 1994). Subject-verb agreement with *be* has, however, rarely been thoroughly investigated in New Englishes. Some scholars carried out diachronic studies on *be* irregularity in Old, Middle and Early Modern English (for example Quirk and Wrenn 1960, Traugott 1972, and Jespersen 1961), others study contemporary English with a diachronic research focus, tracing for example the trajectory of language change with regard to *be* agreement in New Zealand English (Hay and Schreier 2004) or investigating past tense *be* levelling in enclave communities in the mid-Atlantic South, such as Ocracoke, Harkers Island, Smith Island and Mainland Hyde County (cf. Wolfram and Schilling-Estes 2003), while yet other linguists complement this historical evidence by synchronic research of grammatical factors influencing variation in agreement patterns with the verb *be*.

In the native varieties mentioned above, a number of linguistic features have been identified which exert a strong influence on agreement with *be*. These are, among others, tense, the type of subject, polarity, subject number and contractedness. Numerous studies have consistently and repeatedly found the present tense paradigm to be less prone to agreement regularization than the past tense of *be* (c.f. Hay and Schreier 2004: 210). In contrast to this, non-agreement with *be* seems to be literally non-existent in the past tense in Deuber’s 40 direct conversation texts of Jamaican English, and rather infrequent in the present tense. She found 8 cases of the latter, of which one each can be explained by proximity and notional concord (Deuber 2009a) and attributes these instances of present tense *be* non-agreement to the fact that mesolectal creole has invariant *iz* as a copular form for all persons (Deuber 2009a). The subject type is important insofar as existential *there* + *be*-constructions have been found to show higher rates of grammatical non-agreement in past tense contexts with third-person plural subjects than non-existentials, with personal pronoun subjects having the lowest rates (see Hay and Schreier 2004: 211). Some scholars
found an effect of polarity in that *was* is frequently extended to positive contexts, while *weren’t* is often found in negative ones irrespective of the number and person of the subject (Britain 2002). Schilling-Estes and Wolfram (1994: 280) therefore assumed a “remorphologization of the *was* and *were* allomorphs of past *be* along positive / negative, rather than person-number, lines.” However, in Hay and Schreier’s (2004) data there was no evidence of polarity playing any role in determining agreement with *be*. Wolfram and Schilling-Estes (2003) in their investigation of levelling of past tense *be* across several generations in four enclave communities in the mid-Atlantic South found that levelling to *weren’t* has been increasing in negative contexts without, however, a symmetric increase in levelling to *was* in positive contexts. They argue that this can be plausibly explained functionally by the fact that the form *weren’t* “renders negativity more transparent by encoding it not only on the clitic *n’t*, which may be difficult for listeners to perceive, but on the *were*-stem as well,” (Wolfram and Schilling-Estes 2003: 222). They did not, however, find a corresponding functional justification for levelling to *was* in positive contexts. Moreover, the authors point out that “the mainstream educational norm in these communities focuses on the eradication of levelling to *was*” (Wolfram and Schilling-Estes 2003: 224).

Subject number and contractedness are particularly influential factors in a special case of subject-verb agreement with *be*, that is in existential *there + be* constructions (cf. chapter 5 for more details). Such constructions are exceptional in that the controller of agreement is not the element in subject position but the post-verbal noun phrase. Meechan and Foley (1994) found the subject number and contractedness to be particularly important factors determining (non-) agreement in Canadian English existential *there + be* constructions and show that the rate of plural subjects with singular agreement is much higher than that of singular subjects with plural verbs, 72% versus 1% respectively. The authors identify the contracted singular form of *be*, i.e. *there’s* as the most significant indicator of non-agreement and Crawford (2005) even labels this the default form for *there + be* constructions. Crawford, after exploring ETBs (i.e. existential *there + be* constructions) in various spoken and written genres, comes to the conclusion that

> the discourse function of ETBs as topic managers, the tendency of spoken language to use less elaboration and more contraction, combined with the processing constraints of spoken language, results in the formulaic use of contracted existential *there + be (there’s)* without conscious reference to the prescriptive rule of agreement. (Crawford 2005: 59)
The social factors investigated so far as potentially influencing agreement in existential *there + be* constructions include gender, age, social class (see Hay and Schreier 2004 and the summary of previous work given by Walker 2007: 152) and the level of education (Cheshire 1999). Cheshire assumes a close interrelation of educational level and formality as social and situational variables influencing agreement, since she expects “agreement in existential *there* constructions to occur variably: more frequently in formal speech styles than in informal styles, and more frequently in the speech of educated speakers” (Cheshire 1999: 138), i.e. non-agreement would be expected in informal language usage by speakers of lower educational levels. These hypotheses were not confirmed by Crawford (2005) in his investigation of conversations, lectures, textbooks, fiction texts and chat samples in American English. He found existential *there + be* constructions to be more frequent in the academic lectures than in the private conversations and explains this, on the one hand, by referring to the higher cognitive load in lectures and, on the other hand, by the different discourse functions of existential constructions in the text types concerned. Biber et al. point to an influence of the mode of production on agreement in existential *there + be* constructions, as the contracted form *there’s* is often followed by a plural noun phrase in conversation, but not in written language (Biber et al. 1999: 186).

Since, according to Crawford, apart from his work “no study to date has actually compared the extent to which variation is found in a number of different situations of language use” (Crawford 2005: 39), the present paper will focus on register variation along the lines of Biber et al. (1999) and differences between them in the level of formality and mode of production, in order to systematically investigate to what extent these factors influence subject-verb agreement in Jamaican English.
4 Register Variation in Subject-Verb Agreement in Educated Jamaican English

4.1 Introduction

The present study of intra-variety variation aims at investigating the influence of the mode of production and register differences on subject-verb agreement in educated Jamaican English. 'Register' is understood as referring to non-linguistic characteristics of the text type, such as mode of production, degree of interaction, communicative sphere and purpose, and level of formality (see Biber et al. 1999: 15f. for a more detailed account on registers of English). According to Biber et al., such register characteristics of text types are considered to be functional factors influencing linguistic forms: “[L]inguistic features are not uniformly distributed across registers. Rather, there are systematic differences, which correspond to the situational differences among registers (with respect to interactiveness, production circumstances, purpose, etc.)” (Biber et al. 1999: 11). For illustration, Biber et al. refer to the fact that newspaper articles are not interactive but provided by an unacknowledged author to a large and unspecific readership, and this consequently results in the low frequency of first or second person references (1999: 15f.). “However, since newspaper articles are intended to provide current information about important people and events, they commonly use proper nouns referring to known people, places, or institutions” (Biber et al. 1999: 15). That means register differences can be expected to cause different frequencies of particular linguistic forms.

Up to the present, research into New Englishes has focused on features that distinguish one variety from the other, and only few linguists have investigated the way in which functional differences of registers influence the choice of standard and non-standard linguistic forms. It is this lack of research on linguistic variation caused by register differences in New Englishes which prompted the author to carry out the present systematic comparison of registers and their agreement pattern in educated Jamaican English.

The point of departure here is the following question: Is it true that “[p]atwa is clearly the dominant shaping influence on spoken English in Jamaica” as postulated by Mair (2002: 36)? In search of an answer it is necessary to explore the ways in which Jamaican Creole, or Patwa, influences spoken English in Jamaica with regard to subject-verb agreement. For this purpose, three spoken registers differing in their formality levels will be analysed. To
find out whether this postulated influence is really only present in spoken language but not in writing. A parallel analysis of three written registers will be carried out. A subsequent comparison of the analyses of the spoken and written data will show in which registers and to what extent Jamaican Creole exerts an influence on Jamaican English. Moreover, the analysis will elucidate what is the factor that influences non-agreement more strongly, the mode of production, i.e. spoken or written, or the register, represented here by the relative levels of formality, since it will reveal in which of the data categories differences in rates and types of non-agreement are more pronounced and why. The next section provides an overview of the material used for the analysis of register variation in Jamaican English and the methods of analysing the data.

4.2 Data and Methodology

4.2.1 Data – The International Corpus of English

Corpus investigations of World Englishes were greatly facilitated with the advent of the International Corpus of English, a huge project which was designed by Sidney Greenbaum in 1990 to eventually consist of at least eighteen component corpora of regional (and national) varieties of English. The main objective of the project was to overcome the earlier lack of suitable data for contrastive analyses of Englishes – particularly contact varieties of English – around the world by providing linguists with the necessary resources for studying each regional variety of English separately, but also for comparing these eighteen varieties of English around the world.

The corpus comprises material from countries where English is spoken as a native language (ENL) by the majority of the population, e.g. Australia, Great Britain and New Zealand (i.e. so-called ‘inner-circle countries’ in Kachru’s (1997) terminology), but included are also those regions where English functions as an official additional language (ESL), e.g. India and Singapore, amongst others (i.e. ‘outer-circle countries’). In these countries English is generally not spoken as a native language but is used as the only or

41 See Greenbaum (1996a and b) for a detailed description of the project.
42 The varieties that were initially intended to be covered were Australia, Cameroon, Canada, Caribbean, i.e. Jamaica, East Africa (including Kenya and Tanzania), Fiji, Ghana, Great Britain, Hong Kong, India, Ireland (including Eire and Northern Ireland), New Zealand, Nigeria, the Philippines, Sierra Leone, Singapore, South Africa, and the USA. However, the project is growing, as still more research teams have started working on the national varieties of English in their respective world regions. Recently, work was begun on English in Malta, Pakistan and Trinidad and Tobago (For more information see the website of the International Corpus of English project at http://www.ucl.ac.uk/english-usage/ice/index.htm).
major language for government administration, higher education, the media, etc. (cf. Greenbaum 1996b: 4). Not included in the project are countries where English is spoken as a foreign language (EFL) for use as lingua franca in intercultural communication, but where it is not used as a regular means of communication among the natives of the country (Greenbaum 1996b: 4f.), i.e. the countries in the ‘expanding circle’ in Kachru’s (1997) model. Up to the present eight of these one million word core corpora have been made widely available today, those from East Africa, Great Britain, Hong Kong, New Zealand, the Philippines, India, Singapore, and, very recently, Jamaica. The latter will be used as the data basis for the analysis of subject-verb agreement in Jamaican English.

Greenbaum expresses his ideas of possible research based on the ICE-resources as follows:

I envisage the search for typologies of national varieties of English: first-language versus second-language English, British-type versus American-type English, African versus Asian English, East African versus West African English. Researchers might explore what is common to English in all countries where it is used for internal communication, demonstrating how far it is legitimate to speak of a common core for English or of an international written standard. (Greenbaum 1996b: 10)

Since the release of the first corpora, linguists all over the world have been enthusiastically pursuing this goal, using the parallel corpora for (comparative) studies of varieties of English on all continents, for instance Nelson (2006), who compares six of the ICE-corpora in order to determine “the core and periphery of world Englishes.” A large number of studies have until now been published on World Englishes and New Englishes, such as Kachru and Nelson (2006) and Kam-Mei (2000) on Asian varieties of English, Bambgbose, Banjo and Thomas (1995) on West African Englishes, Banjo (1996) on Nigerian English, Kachru (1998) on Indian and Singapore English, Taylor (2001) on the interaction of Australian English with other Englishes, Sand (1999) on Jamaican English, amongst others. All the sub-corpora included in the project comprise 300 spoken and 200 written texts of about 2000 words each, and have roughly the same structure with regard to text categories. The corpora are not completely identical in this respect, however, because in some

---

43 As Greenbaum (1996b: 6f.) describes, the participating research groups are free to expand their respective core corpus of one million words by proportionally increasing the size of all text categories, thereby producing an expanded corpus, by compiling a specialized corpus for particular text categories, by providing non-standard corpora or compiling a continuously updated monitor corpus. For the present study, however, only the core corpora are of interest.

countries it was extremely challenging or totally impossible for the compilers to obtain samples from particular text categories, so that these had to be replaced by the closest equivalent (see Greenbaum 1996b: 5). The corpora may also differ with regard to topics, participant constellations and relationships between interlocutors (see Greenbaum 1996b: 5), as

[it is unreasonable to expect compilers of the corpora to match speakers or writers exactly in the whole range of bibliographical features, such as sex, age, educational level, occupation, or to replicate the types of relationships between speakers in conversations in each corpus in exactly the same proportions. (Greenbaum 1996b: 5)

An essential criterion for inclusion into the corpus is that speakers were born in the country of origin of the respective component corpus or have lived there most of their lives, and that they are 18 years or older. However, effects of age on linguistic variation, e.g. in subject-verb agreement, cannot be studied systematically, because age information is not given for all speakers. The International Corpus of English also does not allow investigation into the influence of social factors such as educational level or social class on subject-verb agreement but can be used only for investigations of educated usage, since only adults are included who “have received formal education through the medium of English to the completion of secondary school” (Greenbaum 1996b: 6). Occasionally, however, the compilers “also include some who do not meet the education criterion if their public status (for example, as politicians, broadcasters, or writers) makes their inclusion appropriate” (Greenbaum 1996b: 6). For this reason, social factors will not be taken into account in the study of subject-verb agreement in educated Jamaican English.

4.2.2 Register-Choice for the Study of Intra-Corpus Variation

The Jamaican component of the International Corpus of English serves as the data for the analysis of register variation in educated Jamaican English. In order to obtain a balanced set of data for both spoken and written language, the text categories need to be carefully selected so as to ensure comparability with regard to their level of formality, and, if possible, also with regard to other characteristics such as the size of the audience, public versus private sphere of communication, etc. The difficulty lies in the fact that the text

45 See Appendix 1 for the text categories in ICE-Jamaica.
categories differ drastically and unsystematically in their register characteristics. While spoken texts can be classified as dialogues or monologues, this distinction does not exist in the written corpus section. Similarly, spoken texts cannot be classified as printed or nonprinted. After careful consideration, the following text categories were chosen as the most suitable basis for the present study, as shown in Table 2:

<table>
<thead>
<tr>
<th>Formality</th>
<th>SPOKEN</th>
<th>WRITTEN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Text category</td>
<td>Text code</td>
</tr>
<tr>
<td>Low</td>
<td>Direct Conversations</td>
<td>S1A-046-065</td>
</tr>
<tr>
<td>Medium</td>
<td>Class Lessons</td>
<td>S1B-001-020</td>
</tr>
<tr>
<td>High</td>
<td>Broadcast News</td>
<td>S2B-001-020</td>
</tr>
<tr>
<td>Total texts</td>
<td></td>
<td>60</td>
</tr>
<tr>
<td>Total words</td>
<td>128,562</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Spoken and written text categories included

From the spoken section of ICE-Jamaica, twenty texts each were taken from the text categories of ‘direct conversations,’ ‘class lessons’ and ‘broadcast news’ to represent three macro-registers differing in their respective levels of formality. The class of ‘direct conversations’ is the most informal text type used, ‘class lessons’ are taken to be of medium formality and ‘broadcast news’ represent the highest formality level. However, these formality levels are not to be taken as absolute, they are relative in comparison to the other text types under consideration, e.g. while class lessons are relatively high in formality when compared to direct conversations, the degree of formality is lower than in broadcast news.

The respective written equivalents are fifteen ‘social letters’ as the most informal text type, twenty ‘student essays’ as intermediate, and twenty ‘press news reports’ as the category of highest formality. While the three spoken categories of twenty texts each, i.e. sixty spoken texts, provide a text basis of around 129,000 words, the analyses of the written material are based on a total of approximately 112,000 words from the three categories. Since all frequencies are normalized, this difference in the size of the data sets does not have a negative impact on the comparability of the results.

Unfortunately, there are only 15 texts in this text category in the corpus. Instead of including five additional texts from another category, and therefore introducing the possibility of skewed results, the decision has been made to leave this category untouched and instead normalize all frequencies obtained.
For the present study, the categories ‘direct conversations’ and ‘social letters’ are assumed to be comparable in terms of their relative level of formality, and so are ‘class lessons’ and ‘student essays’ as well as ‘broadcast news’ and ‘press news reports.’ The text types ‘direct conversations,’ i.e. spoken and informal, and ‘press news reports,’ i.e. written and formal, constitute the two extreme points in the range of stylistic variability covered here, as previous research consistently found that the linguistic variants in written texts tend to be closer to standard English than those in spoken language, and the same is true for highly formal as opposed to informal texts. All other text types selected range somewhere in between these two extremes.

As mentioned above, the registers of each mode of production, i.e. spoken or written, differ not only in terms of formality, but also with regard to other factors such as the size of the audience, the participant constellation, i.e. monologue or dialogue, and the relation of the speakers and audience toward one another, as well as the sphere of communication where the interactions take place, i.e. private or public. However, most of the registers are not homogeneous but are composed of a number of sub-registers, which may differ considerably in the above-mentioned aspects and comprise a wide range of different speech genres. The register characterization below is therefore to be understood not as absolute and true for each and every sample, but as tendencies in relation to the other categories under consideration.

**Registers of Low Formality**

The texts in the categories ‘direct conversations’ and ‘social letters’ are mainly produced in private spheres of communication. Casual conversations are face-to-face interactions involving two or more participants who are present in the immediate communicative situation, know each other well and are roughly equal in terms of social status. They are classified as dialogues. Normally, there is no large audience present and all participants contribute to the discourse in equal proportions, as shown in (1).

\[
(1) \text{<$A><#>He won <#>You go boy you go b <#>And the likl one that was a always smiling <#>It win<giggling</O> <#><{><[>I feel good for him</[</>47 <#>Him can go on and do other things with him life now
\]

\[47\text{The annotation tags }<}|><| …</[ indicate that this stretch of speech is overlapped by another utterance, namely by speaker B in line 4. The first of several simultaneous utterances is indicated by <|>, all utterances before the closing tag </[ are overlapping. Normalization markup of repetitions, false starts, misspelled or incomplete words has been removed from the examples for clarity. Moreover, the transcriptions do not...}
Here, there are three participants watching a spelling competition on TV and commenting on the participants. What is obvious in the example is the high number of overlaps and repetitions, which indicate that the speakers build on previous contributions and jointly produce the discourse. The interlocutors seem to be friends, or at least know each other well, and the general atmosphere of this conversation is very relaxed and jocular, as there is giggling and laughter (lines 2 and 5). This impression is supported by the frequency of discourse markers, such as oh (lines 4 and 7), I mean (lines 9 and 14/15), you know (lines 11 and 14) and anyways (line 15). Moreover, a large number of non-standard features, possibly caused by influence from the creole, can be found in this short sample, such as the use of third person verbs without number or tense marking, e.g. it win (line 2) or him never win (lines 13/14), or the use of him as subject pronoun (lines 2 and 13) or possessive pronoun (lines 3 and 14). All these features indicate that example (1) is a highly informal stretch of conversation between socially close participants.

There is, however, a lot of intertextual variation in the category of ‘direct conversations,’ as not all of the texts are of such extreme interactive and casual type. There are also conversations of a more interview-like nature, where the participants might not know each other well, or where one participant might be clearly dominating the conversation by asking questions or choosing the topics, as speaker A does in (2):

(2)  <$A>$</#>How long have you been working  
<$B>$</#>I’d say about a year and three months now  
<$A>$</#>Was it a planned job or $<$?$$>$like?’the field you wanted to go in  
<$B>$</#>Well actually no because my first degree was in economics but then after

---

contains sentence-final punctuation, such as full stops, question marks or exclamation marks. Instead, beginning text units are marked by $<$#$. See footnote 48 on the notion of ‘sentences’ in spoken language.
5 finishing I decided I didn’t want to go to that area and now I’m doing my masters in management information systems so I think I want to go into the computer-related field (S1A-058)

Speaker A clearly dominates the conversation by eliciting responses to her questions from speaker B and steering the topic of the conversation. There are fewer non-standard features, but the example nevertheless shows characteristics of casual spoken conversation, such as contracted forms (e.g. lines 2 and 5).

In any case, what the above-mentioned spoken examples have in common is that the utterances are produced on the spot, as there is no time to plan grammatical structure or choose words well ahead of time (cf. Biber et al. 1999: 8ff). For this reason, direct conversations are characterised by high numbers of hesitations, repetitions or repair sequences. Participants frequently interrupt one another or speak simultaneously, and sentences are often left incomplete and tend to have few embedded sub-clauses or complex syntactic structures.

Social letters and email messages, such as (3), are similar to conversations in that they are produced for private communicative purposes and the recipients are usually friends or family members of the author, i.e. people who are well acquainted and socially close:

(3) Hi Lorna
    How goes
    I don’t want to sound as if I’m bothering you but could you please tell Peter to do something about that phone either answer it in the days or plug it out because every time I call it just ring for a while and then go busy again.
    Ask him if he can’t ask Diana to allow him to send his messages.
    You take care of yourself and be good. My graduation is on the 9th of November I am suppose to wear white dress and black shoes. Don’t know if I’ll find anything here to suit me but I'll work on that.
    10 Say hi to Phillip and how is the little one.
(W1B-004)

That writer and recipient of this email know each other is obvious, as the recipient is addressed by her first name, Lorna. In addition, reference to other people is also made by using first names, which suggests that they both know these persons, e.g. Peter, Diana and Phillip, or even by phrases which request a high level of familiarity in order to identify the correct reference, namely the little one (line 10).

48 It is questionable whether text units in spoken conversation can actually be called ‘sentences,’ because often the units remain incomplete or important sentential constituents are left out. For instance, there are numerous sentences with zero subject or missing auxiliary verbs in the conversation texts of ICE-Jamaica. However, the term ‘sentence’ is here used to denote the text units in spoken language use, which are usually signalled by means of intonation and often preceded and followed by a pause.
In both text types, ‘direct conversations’ and ‘social letters,’ there is a strong interpersonal focus and topics often centre on private matters, such as common friends or acquaintances and personal relationships, narrations of past events, personal attitudes and opinions, as well as spare-time activities. In example (1) above, the speakers express their opinions on the TV show they are watching together, while in (3) the writer complains that her phone calls remain unanswered. Due to their strong interactional and interpersonal focus, both text types are characterized by high frequencies of verbs in present tense, modals, direct questions, such as the non-standard How goes (3, line 2), imperatives, such as Ask him if…(line 6), You take care of yourself and be good (line 7) and Say hi to Phillip (line 10), and first and second person pronouns (cf. Biber et al. 1999: 8ff. on the text type ‘direct conversations’). In both categories, there is also a large number of contracted forms, such as it’s (1, line 13), that’s (1, lines 8, 11, 12 and 17), don’t (3, lines 3 and 8), I’ll (3, lines 8 and 9), and others.

The text type ‘social letters’ is also characterized by some intertextual variation. Apart from very informal and casual examples like (3), there are also slightly more formal ones, for instance those messages which are sent to persons who are socially not as close as friends or family members, as in (4), which is an email by a student to his or her university professor:

(4) Hi Sir,
I wanted to come and see you today. It’s pouring down outside as I write so I will not be able to leave.
I will definitely come on Wednesday. If I cannot fit into your schedule, I'll leave what I've got for you to criticise. If you find some time, then you can reply to my letter.
Thanks,
<@>name</@>49
(W1B-001)

Even though the focus is not on personal information and there is some social distance between author and recipient, expressed by the address form Sir, the example nevertheless shows many of the characteristics of informal language use, for instance the high frequency of first person pronouns and contractions, as well as the casual greeting formula Hi.

Generally, texts in the category ‘social letters’ tend to display a large number of the non-standard features assumed to be typical of direct conversations,50 such as the lack of

49 Names have been changed or deleted if informants chose to remain anonymous. Such changes are marked with the <@>name</@> tag.
number inflection on third person singular verbs, as in *it just ring for a while and then go busy* (3, line 5) and the absence of –ed marking on past tense verbs or participles, e.g. *I am suppose to* (3, lines 7/8). The sentences tend to be short and of little complexity and information density, but incomplete units are not as frequent as in conversations. In contrast to ‘direct conversations,’ where signs of online production are extremely frequent, hesitation markers, such as *uhm*, overlaps and interruptions do not occur in letters, as there are no other participants present in the immediate communicative surrounding and strategies commonly used in oral communication for keeping or gaining the floor are unnecessary in writing. There are, however, occasional instances of repetitions or self-repairs, for instance when misspelled words are corrected. This indicates that such email messages and social letters are more or less produced on the spot rather than being carefully planned ahead of time.

To sum up, the text categories ‘direct conversations’ and ‘social letters’ are similar with regard to the formality level, the social status of the participants, the sphere of communication and the high frequency of non-standard linguistic forms. Both have a strong interpersonal orientation that is reflected in the topics.

What differs between these categories is that all participants are present in the immediate communicative situation of conversations, but not in letters, where producer and recipient are temporally and spatially distant. Moreover, spoken conversations display a large number of hesitation markers, overlaps and interruptions, i.e. features typical for online produced oral communication, while these are completely absent from the written category ‘social letters.’

**Registers of Medium Formality**

The spoken text category of ‘class lessons’ comprises texts that range in their characteristics somewhere in between the interactive ‘direct conversations’ and the news monologues discussed in the next section. Some of the class lessons belong to the sub-register of lectures and can be described as monologues. In these cases, the teacher as speaker is the only ‘active’ participant and the students assume the more ‘passive’ role of listeners.\(^{51}\) Apart from occasional questions or interruptive comments, they do not

---

\(^{50}\) The validity of this assumption will be tested in section 4.4.

\(^{51}\) According to Bakhtin (2006: 100), the distinction of ‘active’ and ‘passive’ communicative roles is somewhat misleading, as even listeners are actively engaged in taking “a responsive attitude” (2006: 100) to messages conveyed by the speakers. While I fully agree with this statement, I nevertheless use these terms,
contribute verbally to the communication process. On the other hand, there are also a
number of class lessons where teacher and students together discuss a certain topic very
actively. The students here contribute considerably to the utterances made, so that these
examples can be described as highly interactive sequences, where the proportions of
utterances are almost equally distributed between the teacher and the students.
In the sub-register of lectures, where the speaker is not in danger of losing the floor if she
or he needs time for thinking and formulating the utterances, there are comparably few
signs of online-production, while the highly interactive and often fast-paced discussion
sequences are characterized by frequent overlaps or interruptions, repetitions, false-starts
and repair sequences, since here speakers cannot plan ahead but have to compete to take or
keep the floor. However, on average the text category ‘class lessons’ shows fewer of these
features than the informal text type ‘direct conversations.’
Such by-products of unplanned discourse are largely absent from the written texts of the
same formality level, the category of ‘student essays,’ which comprises non-printed and
unpublished timed examination essays written for one or two examiners whose identity
may or may not be known to the writer, and untimed student essays written to be read and
evaluated by the lecturer. Indications of editing are sometimes found in both sub-registers,
e.g. when misspelled words are crossed out and corrected, but they are not frequent. For
illustration, an example of a correction\textsuperscript{52} within a student essay is given as (5).
\begin{quote}
(5) Our economy at this time cannot afford free tertiary education, \texttt{but} \texttt{so}
we must discover new ways to fund this venture. The \texttt{ambit} \texttt{scope} of
the Students Loan Bureau may have to be reorganized, but ways must be found to
make the university more accessible to the ordinary Jamaican.
(W1A-001)
\end{quote}
The topics covered by the essays are not as diverse as those in the class lessons, as quite a
number of essays by different students were included in the corpus which deal with the
same topic. In class lessons and essays, both the presentation of information and the
expression of personal opinions are important communicative purposes. However, the texts
of medium formality are more concerned with knowledge and information than personal
topics, relationships and spare-time activities and are thus more impersonal than the text
types of low formality. In other words, with respect to the topics these categories range

\textsuperscript{52} Items that are crossed out from the student essay are marked by \texttt{…}
Usually the following word represents the correction of that misspelled or badly chosen word. Here, both are highlighted in bold.
between ‘direct conversations’ and ‘social letters’ on the one hand and oral and written news texts on the other. Observations of linguistic features show that in ‘class lessons’ and ‘student essays,’ there tend to be considerably fewer non-standard features than in the most informal text types, but more than in the highly formal news texts. Syntactic structures in student essays tend to be more complex than in informal letters, and the former are also much more argumentative.

**Registers of High Formality**

Last but not least, the selected registers of highest formality are ‘broadcast news’ and ‘press news reports.’ ‘Broadcast news’ are typically highly formal monologues in the public sphere, in which one speaker, who is not necessarily the originator of the text, communicates something to a large unknown audience that is not present in the immediate communicative surrounding. However, these monologues are sometimes interrupted by reports by other speakers, whether they are correspondents or interview partners, but only in very rare cases does that lead to a more interactive communicative style. Example (6) is an extract from one of the pure monologues in the corpus:

(6)  

A very pleasant afternoon to you  
Here now is the latest news  
R J R news understands that the Bank of Jamaica has identified consistent violations of certain banking regulations by the management of Tower Merchant Bank and Trust Limited which has led to strong action being taken by the Finance Ministry  
This morning the national news agency Jampress reports that Finance Minister Hugh Small has taken control of the management of the institution as of today (S2C-001)

Since broadcast news are pre-formulated, i.e. ‘scripted’ (see Nelson 1996: 31) and read out to the audience, there are generally very few signs of online production in this category, such as repetitions, incomplete utterances, hesitations or false starts, and the grammatical information tends to be encoded more synthetically than in most other text types, such as ‘direct conversations’ or ‘class lessons’ (Szmrecsanyi and Kortmann 2009). Due to the factual orientation to past events, third person reference is particularly frequent in news texts (cf. Biber et al. 1999: 11), while direct questions, imperatives, or first and second person pronouns are extremely rare in this non-interactive type of communication. The purpose of news texts is to present factual information about past events objectively and effectively, i.e. to communicate the content in very little time. This objective causes news presenters to use structurally complex sentences (cf. Nelson 1996: 32) with a high
information density and high proportions of proper nouns, passive voice (6, line 4), verbs in past tense, participle constructions and prepositional phrases (lines 2/3 and 6), but few direct questions or contracted forms (cf. Biber et al. 1999: 8ff.).

Similar observations can be made for the texts in the category of ‘press news reports,’ an example of which is given as (7).

(7) Charges of fraud at JADCO
DIRECTOR of Public Prosecutions, Glen Andrade, and Auditor-General Adrian Strachan, have, in their investigations into the irregularities of the accounts of the Jamaica Attraction Development Ltd. (JADCO) found that fraud appeared to have been committed.
This disclosure was made at yesterday's sitting of the Public Accounts Committee when a letter from Mr. Andrade and a Special Report on JADCO from the Auditor-General were made public.
(W2C-011)

These highly formal written texts are produced with the purpose of objective presentation of information in the public sphere by an unspecified author for a large unknown readership, and author and addressees do not share the immediate communicative context. There is little personal opinion conveyed in these texts, but the texts are written on behalf of an institution, namely the respective news agency, as shown in (6, lines 1/2 and 5). ‘Press news reports’ and ‘broadcast news’ are similar in most respects, for instance, they undergo a process of careful planning, elaboration and editing before being published and made available to the intended addressees; both are impersonal and institutional rather than personal and interactive, and both have an informational focus and are structurally complex with passive constructions (7, lines 4/5, 6 and 7/8) and embedded clauses (7, lines 3/4), to name but two features. Contracted or non-standard forms are rare in both text types. The main difference is the mode of production. While press texts are written texts that are printed, broadcast news are pre-formulated but spoken monologues.

Summary
Table 3 shows a summary of the most central distinguishing characteristics of the registers53 selected for analysis:

53 In the table, the text categories are abbreviated as follows: ‘direct conversations’ = CONV, ‘social letters’ = MAIL, ‘class lessons’ = CLASS, ‘student essays’ = ESSAY, ‘broadcast news’ = NEWS, and ‘press news reports’ = PRESS.
As mentioned before, the text types of lowest formality are the spoken ‘direct conversations’ and the written ‘social letters,’ those of medium formality consist of the spoken ‘class lessons’ and the written ‘student essays,’ and the most formal texts in the study are the spoken ‘broadcast news’ and the written ‘press news reports.’

In the written section, the primary distinction is that between printed texts, i.e. those that are intended for publication and addressed at a large unknown readership, and non-printed texts, which are not intended for publication and whose readership is much smaller and often personally known to the author (cf. Nelson 1996: 31f.). The most informal written text type under consideration, i.e. ‘social letters,’ and the category of medium formality, i.e. ‘student essays’ are non-printed, while the highly formal ‘press news reports’ are printed. In spoken texts the main distinction is between dialogues, involving two or more participants, and monologues produced by a single speaker. While ‘direct conversations’ are highly interactive dialogues in which the participants share the immediate communicative situation, the text type ‘class lessons’ comprises both interactive dialogues and non-interactive monologues, so that one can speak of ‘restricted interactiveness’ in this category. Here, there is also co-presence of the participants in the communicative context. The formal ‘broadcast news’ are almost exclusively monologues and thus not interactive. Producer and addressees are temporarily and physically distant, i.e. there is no shared communicative situation.

Indications of online planning, for instance hesitation markers, repetitions, false starts, and so on, can be found mainly in spoken texts of low and medium formality, but also, albeit rarely, in low and medium formality written texts. They are absent from both ‘broadcast news’ and ‘press news reports,’ which are carefully planned and edited before being
published. Only the text categories of lowest formality, i.e. ‘direct conversations’ and ‘social letters’ are aimed at personal communication, while all others are intended to be informative. The text types of lower and medium formality are all addressed to individual recipients, for example friends, colleagues or family members in the case of conversations or letters, and student essays are usually produced in order to be read by the teacher or examiner. Only news texts, whether written or spoken, are addressed to a large unknown, i.e. public, audience.

4.2.3 Methodology – Analysis of the Registers

In a first step, subject-verb agreement with lexical verbs and with the verb be is analysed in the spoken and written text categories of ICE-Jamaica separately, in order to find patterns of register variation based on the relative levels of formality. After that, the results of each mode of production are compared and emerging patterns discussed.

For the analysis of subject-verb agreement with lexical verbs I manually\textsuperscript{54} counted all verbs in present tense with third person singular or plural subjects except modal auxiliaries and be. I excluded from my counts all sentences with null subjects (8) or deleted copular or auxiliary verbs (9), as well as subjunctive or tag questions.

\begin{align*}
\text{(8)} & \quad \text{Seems like she’s funny (S1A-056)} \\
\text{(9)} & \quad \text{And a lot of times because they’re sitting up there they not see what’s happening on the ground in terms of the students (S1A-059)}
\end{align*}

Moreover, questions were not included, as it was often impossible to determine whether the example under consideration was to be counted as grammatical non-agreement, or whether the auxiliary verb was missing and the example was not to be counted at all. An example of such a doubtful case is the question given in (10).

\begin{align*}
\text{(10)} & \quad \text{So that mean you wouldn’t watch uhm <quote>Dawn of the Dead</quote> (S1A-051)}
\end{align*}

\textsuperscript{54} Manual analysis of the data is necessary, as ICE-Jamaica is not tagged for part of speech and the context thus has to be taken into account for the identification of unmarked plural nouns or third person singular verbs.
In this utterance, there is either grammatical non-agreement between the subject pronoun *that* and the verb *mean*, or the auxiliary *does* is missing. Since counting these cases would have been rather arbitrary, the decision was made to exclude all questions.

Occasional unclear cases do not appear in any of the analyses, either. These are most frequent in the direct conversations, and rarest in the broadcast news. The three most frequent reasons for exclusion are unclear subject (11), unclear tense (12), where there is neither marking for past tense nor number, and sentences which are not clearly interpretable or show clusters of creole features (13).

(11) This in an effort to *improve* cooperation and communication in areas which affect the whole of Ireland (S2B-007)

(12) Something *drop* (S1A-050)

(13) … they start cuss and how they spent ten thousand dollars on campaigning and them never win is for me *why* me win (S1A-060)

Since the controller of agreement is normally the head noun of the subject noun phrase, the decision has been made to focus on overt marking of nouns and to count as non-agreement those instances where there is non-agreement within the subject noun phrase (14). That is, the head of the subject noun phrase is not marked for number but the plural is indicated by other means, such as quantifying expressions:

(14) Oh man alright so uhmm that's uh why I *oh* yeah yeah get in trouble and all sort of thing happen (S1A-050)

These cases are, however, quite rare. Contributions by extra-corpus speakers, quotations and creole passages indicated by <X>…</X>, <quote>…</quote> and <indig>…</indig> tags respectively, were also not taken into consideration. Since ICE-Jamaica is tagged for a number of discourse phenomena, these had to be taken into account as well, so that instances marked as unclear speech <unclear>…</unclear> or uncertain transcription <?>…</?>, but also self-corrections and repetitions <->…</-> were excluded from the analyses.

---

55 Extra-corpus speakers are those speakers who are not Jamaicans, for instance linguists from the Freiburg research team who collected the material and sometimes participated to a certain extent in the conversations, or people who do not qualify for inclusion into the corpus on other grounds, e.g. educational specifications.
Cases such as (15), where the word following the verb begins with a sibilant, had to be disregarded as well, because the pronunciation of these words makes it impossible to identify whether or not the verb is singular, as in (15a), or plural, as in (15b).

(15a) I know that there's some negatives to that because they are sometime not uh m
(S1B-018)

(15b) I know that there’re some negatives to that…

Due to this difficulty in determining the exact form of the verb, these instances of be occurring before sibilants are excluded from the analysis.

The point of departure for the categorization of the verbs and their inflection is grammatical agreement. This means that agreement marking that does not correspond to the grammatical number of the head noun but is allocated according to the semantic notion of the subject is marked as non-agreement, which, however, is not to be taken as an evaluation of the correctness of the form. It merely expresses that there is no grammatical agreement.

For the analysis of be I looked at present and past tense verbs following subjects in all persons and numbers. Non-agreement was almost non-existent in all persons except third. There was one single instance of is occurring with a second person singular subject in the text category ‘class lessons.’ This is rather surprising because, as Deuber (2009a) notes, is is used in Jamaican Creole as an invariant copular form with all persons. For this reason a considerable number of such instances were expected to occur especially in direct conversations, but no such cases were found in my data. Since non-agreement occurred almost exclusively with third person subjects, I will restrict my analysis to instances of be agreeing (or not agreeing) with third person subjects. Excluded from my counts were the same cases as above, namely sentences with zero subjects or absent copular verb, questions, tag questions and the annotation tags mentioned, as well as very few instances where the subject was not clear.
4.3 Results for Spoken Jamaican English

4.3.1 Agreement with Lexical Verbs

A total of 4088 verbs were included in the analysis of lexical verbs, of which 2055 verbs were found in the spoken and 2033 verbs in the written corpus section. Table 4 shows the distribution of these lexical verbs across the text categories. The three text types in the spoken section are ‘direct conversation,’ ‘class lessons’ and broadcast news,’ which are here represented by low, medium and high formality level, respectively, and the corresponding written text types are ‘social letters,’ ‘student essays’ and ‘press news reports.’

<table>
<thead>
<tr>
<th>Mode of Production</th>
<th>Spoken</th>
<th>Written</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tokens</td>
<td>%</td>
<td>Tokens</td>
</tr>
<tr>
<td>Low</td>
<td>713</td>
<td>65.5</td>
<td>376</td>
</tr>
<tr>
<td>Medium</td>
<td>638</td>
<td>38.2</td>
<td>1030</td>
</tr>
<tr>
<td>High</td>
<td>704</td>
<td>52.9</td>
<td>627</td>
</tr>
</tbody>
</table>

Table 4: Distribution of lexical verbs across the spoken and written text categories

Of the 1089 verbs of low formality, 65.5% occurred in the spoken and 34.5% in the written texts. The distribution is reversed for the categories of medium formality, where 61.8% of the 1668 verbs occurred in written and only 38.2% in spoken texts. The verbs in higher formality texts are more evenly distributed, with 52.9% in spoken and 47.1% in written texts. For the time being, the mode of production is disregarded and only the spoken registers are analysed with regard to agreement and non-agreement. Written language use will be discussed in detail in section 4.4.

The 713 verbs in the category ‘direct conversations’ represent 34.7% of all spoken lexical verbs, the 638 verbs in ‘class lessons’ add up to 31% and the 704 verbs in the category ‘broadcast news’ constitute 34.3% of verbs other than be. The lexical verbs are thus distributed fairly evenly across the three spoken registers. Table 5 gives the token numbers

---

56 An earlier version of this chapter is forthcoming as “Agreement in educated Jamaican English: A corpus-based study of spoken usage in ICE-Jamaica.” In Anja Wanner and Heidrun Dorgeloh. Approaches to Syntactic Variation and Genre. Berlin: Mouton de Gruyter. Some of the categories have been slightly refined, however. For this reason, the numbers and results reported here may differ slightly from those in the paper.

57 The verbs have and do are included under lexical verbs in order to separate them from the exceptional verb be. Their agreement patterns are analysed irrespective of whether they are used as auxiliary verbs or full verbs.

58 The verbs occurring with the exceptional noun police are not included in this count, as these will be analysed separately.
and percentages of grammatical agreement and non-agreement in the spoken categories of low, medium and high formality respectively.

<table>
<thead>
<tr>
<th></th>
<th>Grammatical Agreement</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Agree</td>
<td>Non-agr</td>
<td>Total</td>
</tr>
<tr>
<td></td>
<td>Tokens</td>
<td>%</td>
<td>Tokens</td>
</tr>
<tr>
<td>Formality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>588</td>
<td>82,5%</td>
<td>125</td>
</tr>
<tr>
<td>Medium</td>
<td>586</td>
<td>91,8%</td>
<td>52</td>
</tr>
<tr>
<td>High</td>
<td>679</td>
<td>96,4%</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>1853</td>
<td>90,2%</td>
<td>202</td>
</tr>
</tbody>
</table>

Table 5: Distribution of agreement and non-agreement with lexical verbs in the spoken registers

The overall rate in all 60 spoken texts is 9.8% for non-agreement and 90.2% for agreement. This means that non-agreement is not a frequent phenomenon. However, looking separately at each level of formality, some systematic differences between the categories come to the fore in that non-agreement rates decrease steadily with increasing levels of formality, namely from 17.5% in low formality, to 8.2% in medium formality, and then down to 3.6% in high formality. A Chi square test of significance shows that the influence of the level of formality on the rate of non-agreement in the spoken data is highly significant at the 0.001 level.59

Figure 3 shows the percentages of grammatical non-agreement by subject number for the three registers, i.e. levels of formality. Here, agreement rates are displayed separately for verbs with singular as opposed to plural subjects, and a very clear agreement pattern emerges.

59 The Chi square test yields \( \chi^2 = 81.1 \). This clearly exceeds the critical value of 13.82 at the 0.001 level for 2 degrees of freedom (df).
Figure 3: Grammatical non-agreement by subject number (spoken)

In all levels of formality, non-agreement rates are higher when the subject is singular, as in (16), than when the subject is plural (17).

(16) And we find **it help** you know when the spouses are there (S1A-054)
(17) … and you have **people** out there who **wants** to see these things… (S1A-051)

In low formality direct conversations more than one-third, 35.2% of all verbs with singular subjects show a lack of –s marking. This rate decreases with increasing formality to 12.4% in class lessons and then to 3.8% in broadcast news. On the other hand, the rate of non-agreement with plural subjects, which is around 1% in low and medium formality texts, increases in the most formal category, where 3.1% of the verbs following plural-subject noun phrases show hypercorrect third-person singular –s marking. This distribution suggests that the lack of –s marking is stigmatized, as it decreases with increasing formality, while the third-person singular inflection is a prestigious form, which increases with increasing formality. In all text types, non-agreement is much more likely to be characterized by a lack of marking rather than by too much, or hypercorrect, marking. As expected, a chi square test shows that subject number is a highly significant factor determining total rates of grammatical agreement and non-agreement in my spoken data ($\chi^2 = 107.0$ exceeds the critical value of 10.83 at the 0.001 level for df = 1). However, in broadcast news the difference in non-agreement rates between sentences with singular as
opposed to plural subjects is only 0.7%, and this difference is statistically not significant. The reason for this pattern is more likely the steep decrease of non-agreement with singular subjects when the level of formality rises rather than the much less steep increase in hypercorrect marking.

In order to find out whether non-agreement is caused by the same factors in all text types we have to look at what might account for hypercorrect marking of verbs with plural subjects and what might cause the frequent lack of marking with singular subjects, as well as what other types of non-agreement are present in the data. The types of non-agreement in the three spoken text categories are given in Figure 4, where ‘absent –s’ denotes cases of absent third person singular –s marking; ‘Do negative’ represents the simplified, invariant use of the contracted don’t as the negative form with all persons and numbers in the present tense; ‘NP-marking’ means that there is non-agreement within the noun phrase, i.e. the whole NP is intended as plural and marked by means of quantifiers or the context but the head noun lacks the –s marking; ‘Hypercorrect’ denotes verbs following third person plural subjects which nevertheless have singular –s marking. Lastly, the category ‘Other’ includes all those cases which cannot be classified under any of the types mentioned above.

The chi square test of the correlation between subject number and the rate of non-agreement in the category broadcast news yields 0.125. This does not exceed the critical value and is thus not significant. The fact that subject number is a highly significant influence on non-agreement rates in conversations and class lessons but not in broadcast news indicates that not subject number alone determines agreement patterns but that number interacts with another factor, a factor that causes the high non-agreement rates particularly with singular subjects in conversations and, to a lesser degree, in class lessons, but which does not exert any influence on agreement patterns in broadcast news.

See Trudgill (1990) for a discussion of don’t simplification in British dialects. In some British dialects there is a tendency to use don’t as the invariant form of negative do with all persons and numbers, i.e. to extend it to third person singular contexts, making possible forms such as he/she don’t. The verb do will be discussed in more detail below.
The one instance of ‘other’ types of non-agreement in my spoken data is given as (18).

(18) ... even in the face of what appears to be overwhelming odds\textsuperscript{62} (S2B-016)

Despite the controller of agreement being odds, the plural head noun of the post-verbal NP overwhelming odds, the verb appears is singular and thus does not agree grammatically with its subject.

Absent –s marking in third person singular verbs, as in (19) is by far the most common type of non-agreement in all three levels of formality.

(19) I don’t know which way she drive that is a she (S1A-050)

While this type accounts for 81.6\% of all non-agreement occurring in the text category conversations (102 tokens) and 82.7\% in the class lessons (43 tokens), it represents 72.0\% of non-agreement in broadcast news (18 tokens). In terms of total numbers, this type of non-agreement decreases sharply with increasing formality. A similar decrease can be observed for non-agreement involving don’t, the contracted negative form of do, which occurs 16 times in conversations (that is, 12.8\% of the non-agreement in low formality texts), 7 times in the class lessons (13.5\% of the non-agreement in this category) and is completely absent from broadcast news. In (20), an example is given for non-agreement with do negative.\textsuperscript{63}

\textsuperscript{62} Quirk et al. show that “with the determiners what and whatever, the concord depends on the number of the determined noun” (1985: 755). In the example the verb should therefore be plural for grammatical agreement to take place.

\textsuperscript{63} Agreement and non-agreement with the verb do will be analysed separately below.
(20) …I’m saying she’s like she don’t really divorce (S1A-062)

Non-agreement resulting from NP-internal inconsistencies of plural marking, e.g. (21), occurs exclusively in texts of lowest formality and (with 3 instances) accounts for 2.4% of the amount of non-agreement in this text type.

(21) As soon as it the twenty-one day that you’re to take finish then a period will occur (S1A-054)

Last but not least, hypercorrect marking, as in (22), increases with increasing formality from 3.2% in conversations to 24% in broadcast news.

(22) …and you have the hall chairs who also gets money (S1A-059)

In the broadcast news texts, one example of hypercorrect marking may be attributed to the proximity principle (23), according to which the verb sometimes does not agree with the syntactic head but with a nearer noun in a complex noun phrase:

(23) The moves to replace Mr. Sangster follows three successive losses to Mr Buchanan in general elections (S2B-009)

Here, the syntactic head is the moves, which is plural, but the verb has third-person singular –s marking due to the more proximal singular noun Mr. Sangster. None of the hypercorrect markings in the other categories can be characterized as proximity agreement. Three of the hypercorrect verb forms in the broadcast news category may have been marked for singular due to the notional rather than the grammatical number of the subject (24), whereas only one instance in the class lessons category and none in direct conversations category can be so explained.

(24) The ten includes three unidentified men two security guards and a gunman (S2B-011)

In (24), the head of the subject noun phrase is interpreted as “the total of ten” and therefore causes singular agreement. That hypercorrect marking is most frequent in highly formal texts can be attributed to third person –s marking being the more prestigious form. This prestige of –s marking is also reflected in the fact that the noun police, which takes a plural verb in an overwhelming majority of cases in Jamaican English, is occasionally followed
by a verb marked for third-person singular. This, however, can only be found in the highly formal text categories of broadcast news, e.g. (25), and press news reports.64

(25) **The Red Hills Police says** it has received no formal report nor is it aware of a councillor from the area brandishing a firearm before demonstrating residents on Monday (S2B-002)

In (26) the noun police is followed by two verbs, one of which is plural and one is singular.

(26) **The Linsted police** who **are** investigating **has** since arrested Arnof on a charge of murder (W2C-005).

Due to the frequent occurrence of this or similar initial phrases it can be hypothesised that the phrase functions as a kind of formulaic expression in news texts and is not processed as deeply as other subject verb structures. That could explain that police is used once with plural verb and once with singular verb even within the same sentence.

As concerns the lack of inflection on verbs following singular subjects, the strongest factors seem to be strikingly different in each of the categories, as is shown in Figure 5.

---

**Figure 5: Potential explanations of non-agreement involving absence of –s marking across formality levels (spoken)**

---

64 In a total of 61 instances of police + V in the spoken text category of broadcast news two are followed by a singular verb form (that is 3.3%), while in the written press news reports three out of 14 instances are followed by a singular verb (21.4%).
For low formality direct conversations, of the 102 verbs that do not grammatically agree in number with their subjects due to a lack of –s marking, only one case could be described as proximity agreement (27):

(27) There’s <<?>one</?> oh one of my lecturers come from I think not Linstead Sligoville (S1A-050)

Here, not the syntactic head of the noun phrase one, but the closer noun lecturers functions as the controller of agreement. This ties in well with Deuber (2009a: 22), who finds that proximity agreement is extremely rare in direct conversations. In fact, not a single instance of non-agreement in her sample of about 80,000 words can be explained by proximity agreement. In the text categories of class lessons and broadcast news, two lexical verbs each lack third-person singular marking due to proximity agreement. Therefore, one cannot speak of a striking difference between the registers and formality levels with regard to the frequency of proximity agreement with lexical verbs, and proximity is not a crucial factor influencing non-agreement in spoken Jamaican English.

Notional agreement is somewhat more frequent in the most informal text category; 17 instances, or 16.7% of the cases of non-agreement in conversations, can be thus explained. Of these, ten occur in phrases that have an indefinite pronoun as the subject (28), including somebody (six tokens) and everybody (two tokens).

In one case each + noun is interpreted as plural and followed by a plural verb, and the other six instances occur in phrases in which the head noun may be given a collective reading (29).

(28) Everybody have to go to the chapel (S1A-056)

(29) Otherwise the I’ll say that administration have a lot of respect for the Student Guild (S1A-059)

Interestingly, there is variation in the marking of verbs following such collective nouns, in one case even within a single utterance (30).

(30) Hollywood puts two and two together and just give you what you want (S1A-051)

---

65 If only the orthographic transcription is considered, example (27) could be a non-standard zero relative clause, i.e. There’s <<?>one</?> oh one of my lecturers ∅ come from I think not Linstead Sligoville, but from the audio recording it becomes clear that oh is stressed and the following clause is introduced as new information. Therefore, this utterance is most likely not a case of zero relative.
In class lessons, only one instance (2.3%) of grammatical non-agreement can be attributed to notional considerations, and the verb is in this case preceded by an indefinite pronoun. In broadcast news, 11 instances of notional agreement can be found (61.1% of non-agreement), ten of which involve nouns that can be understood as denoting either a group or several individual members of that group, as in (31), and one involving each + noun used with a plural verb.

(31) Despite the uncertainty of the tour taking place the West Indies Cricket selection committee have chosen a sixteen-man squad… (S2B-017)

The frequency pattern of notional agreement seems rather surprising; the numbers are highest in the most informal category of direct conversations, then they decrease sharply and there are hardly any instances in class lessons, and the numbers increase considerably again in broadcast news, the most formal text type. However, when the type of notional agreement is taken into account, it becomes clear that register characteristics are highly influential here. Conversations are the register in which most of the phrases with indefinite pronouns, such as everybody and somebody, occur (81 tokens). These constructions are rather infrequent in class lessons. Of the 56 tokens of *body in class lessons, most occur in questions, as in (32), and thus were excluded from analysis.

(32) Anybody want to help me with it (S1B-011)

In conversations, seven instances can be found where personal pronouns refer back to an indefinite pronoun, as in (33), all of which are plural.

(33) Don’t trust anybody to teach me to swim cos they’re gonna drown me (S1A-052)

The few pronouns in the other text categories (spoken and written) referring back to indefinite pronouns are also plural. This might be taken as an indication that these indefinite pronouns are indeed interpreted as denoting plural number, but it can also point to a tendency to avoid using gender specific pronouns for indefinite subjects.

The pronouns in question are absent from the 20 broadcast news texts. Therefore, the overall distribution of indefinite pronouns correlates negatively with formality levels in the spoken text types. In contrast to this, the number of collective nouns rises with the level of formality. A search of Hundt’s (2006) 35 collective nouns and an additional 7 nouns
(administration, exec(utive), faculty, Hollywood, office, Seacole and UWI) reveals that these are most frequent in broadcast news (518 instances), and almost equally frequent in class lessons and direct conversations (196 and 201 cases, respectively). Therefore, the rise in notional agreement in news texts is not really surprising anymore. As has become clear from the above discussion, the difference in the rate of notional agreement between conversations and class lessons is not due to different distributions of collective nouns but instead to the high number of indefinite pronouns in the former category and the low number of these pronouns in the latter category.

However, not only do the numbers of occurrences differ across the registers, but so too do the types of collective nouns. While direct conversations are particularly rich in collective nouns such as (in order of frequency) family, university, and UWI, the most frequent collective nouns in class lessons are group, class, and government. In broadcast news, ministry, government, and company far outnumber all other collective nouns. These frequencies reflect the characteristic distribution of topics in the text categories discussed and thus provide evidence for the strong influence of formality and register characteristics on the choice of vocabulary and syntactic structures such as subject-verb agreement patterns.

The above analysis of different types of notional agreement has shown that Deuber’s assumption of a tendency in Jamaican English to extend notional agreement can be seen only in low formality conversational texts, where indefinite pronouns are most frequent, but not in broadcast news, where the number of collective nouns is highest. In the former category, indefinite pronouns, such as everybody, which are “unambivalently singular” (Quirk et al. 1985: 764) in International Standard English, are notionally interpreted as plural, thus notional agreement is extended. In the latter register, on the other hand, there are few such reclassifications of nouns. The collective nouns in this category, government for example, can be interpreted even in International Standard English as entities or groups of persons, hence singular in number, or as the individual members of the category, hence plural in number. Moreover, agreement with collective nouns functions differently in different varieties of English. While American English “generally treats singular collective nouns as singular” (Quirk et al. 1985: 758), there is more variation in British English and the number of the verb depends on the interpretation of the collective noun as denoting an

---

66 In the direct conversations, executive was occasionally abbreviated to exec, so a WordSmith search in the three text categories was carried out for EXEC*.

67 UWI is an acronym that stands for the University of the West Indies.
entity or several individual items (Quirk et al. 1985: 758). Jamaican English has variable agreement with collective nouns and thus follows British English more closely. However, although plural agreement is possible with many collective nouns it is not generally the preferred form.

The tendency to extend notional agreement in informal texts is not only visible in the case of indefinite pronouns, but the same is also true for the use of the quantifier much, which can be extended to contexts in which it determines (plural) count nouns (34):

(34) I don’t have **that much loads** (S1A-046)

This phenomenon is more frequent in conversations (where it occurs 56% of the time, or in five cases) than in class lessons (where it occurs 33% of the time, or in four cases, three of which are repetitions of the same sentence *How much times a week* (S1B-010) uttered by the same speaker), and it is completely absent from broadcast news. While the number of occurrences is too low in this case to make reliable assumptions about reclassifications of nouns along the count versus non-count distinction, these numbers do tentatively support Sand’s (1999: 134f.) results.

A few initial observations will help in interpreting the results as we turn from notional considerations to further types of non-agreement. Since casual speech is furthest removed from the standard variety and closest to Jamaican Creole, a wide range of upper and even some lower mesolectal forms can be expected in informal conversation. One of the pervasive features of creole languages, e.g. Jamaican Creole, is the general lack of inflectional morphology (see Patrick 1999). Previous research has shown that the corresponding lack of –s marking in third person singular verbs is a prominent characteristic in the Jamaican mesolect (see for example Sand 1999: 133). The absence of -s marking in Jamaican English is therefore assumed to occur frequently particularly in those environments in which other creole features are evident. By contrast, in highly formal language, such as that commonly used in broadcast news, very few or no creole features are expected, as highly formal language is located closest to the standard on the creole continuum. As an example of such creole features, we will consider the use of the third-person pronoun him in subject position (cf. Patrick 1999: 429 and Deuber 2009a: 28). The data support the above assumption: there are 15 instances of absent –s marking in such mesolectal environments in the register of direct conversations (14.7% of all non-agreement in this text type), while there are no such instances in both class lessons and broadcast news. The fact that such creole pronouns occur mainly in conversations indicates
that their use is highly salient and stigmatized in Jamaican English. An example of non-agreement in a creole context is given in (35).

(35) So **him** just still **go** at that school until now (S1A-061)

There is only one instance in the data where **him** functions as the subject and the verb is nevertheless marked for singular (36).

(36) … and that way **him has** produced a little brother… (S1A-049)

Apart from this one exception, **him** in subject position practically excludes the possibility of the verb being marked for singular. This can be explained by the fact that the use of **him** as a subject pronoun is an overt mesolectal feature, while –s marking on verbs is a feature that is located much closer to the acrolectal end of the creole continuum. The infrequency of the co-occurrence of these features shows that speakers do not choose randomly any variants from their linguistic repertoire but that the variants chosen tend to be located in similar positions of the creole continuum. In other words, if speakers use an acrolectal feature such as singular inflection on verbs they will most likely also use other acrolectal features rather than mesolectal features, and vice versa.

The pronoun **him** can be used not only as a personal pronoun functioning as the subject, but occasionally it is also used as a possessive pronoun (cf. Deuber 2009a: 27f.), as in (37).

(37) **Him** never win and **him brother** win and… (S1A-063)

This example contains an instance of both creole-type uses of **him**, as a personal pronoun in subject position at the beginning of the sentence and, in addition, as a possessive pronoun in the second noun phrase, and thus provides evidence for this tendency to use clusters of several features from roughly the same position on the continuum. As expected, both creole pronouns are in connection with an uninflected verb. The corpus analysis shows, however, that uninflected verbs are common in all three spoken registers.

In all text categories except that of broadcast news, the simple unmarking of verbs for third-person singular is the most frequent type of non-agreement, accounting for 67.6% of all non-agreement cases in direct conversations and 88.4% in class lessons. In the most formal category, unmarking is the second most influential factor, accounting for 27.8% of non-agreement, after notional considerations, which are responsible for 61.1% of non-agreement in this news category. The percentages suggest that unmarking is more frequent in class lessons than in direct conversations. However, this percentage shows only the
proportion of non-agreement explained by simple unmarking as compared to the total amount of non-agreement in the respective category. The absolute number of unmarked verbs declines from 69 instances in texts of least formality to 38 in texts of medium formality and finally to 5 tokens in texts of highest formality, or broadcast news. The finding that unmarked verbs, although decreasing in frequency with increasing formality, do occur in all text categories suggests that the lack of inflection is a much less salient feature than the use of overt creole pronouns. While uninflected verbs are common also in vernacular varieties of English that do not have a creole substrate, its presence and pervasiveness in the Jamaican mesolect and its high frequency particularly in informal Jamaican English proves that the creole exerts an indirect influence by facilitating its occurrence in Jamaican speech.

Last but not least, we will consider what is called ‘Do negative’ in Figure 4, i.e. the invariant use of the contracted *don’t* as the negative form with all persons and numbers in the present tense. Invariant *don’t* has been found to be pervasive in some British dialects where it serves to simplify the inflection system (cf. Trudgill 1990 for a discussion of *don’t* simplification in British dialects). The verb *do* was found 160 times in my spoken data, in 82.5% (132 tokens) agreeing and in 17.5% (28 tokens) not agreeing with their subject. The majority of non-agreements occurred in sentences where the subject was singular, while only two non-agreeing verbs followed plural subjects.

In view of this unequal distribution, only non-agreement with singular subjects will be considered further below. Figure 6 shows the distribution of agreement and non-agreement with the verb *do* in positive versus negative and in full versus contracted forms for the three text categories representing low, medium and high levels of formality.

---

68 Only present tense occurrences of *do* are relevant for the discussion of number marking. Other uses of the verb are not considered here.
Here ‘Do’ means that the form is positive, as in (38).

(38) … yes that’s what she do language and communication I think (S1A-050)

Due to the low number of tokens no distinction was made in the analysis between cases in which do is used as a full verb as opposed to an auxiliary verb. However, uses as a full verb are completely absent from all but the first usage type, that is, they occur exclusively in positive contexts. In the vast majority of cases do is used as an auxiliary verb.

Cases in which the sentence is negated but the full form is used are denoted ‘Not do,’ for example (39), and all contracted negative forms appear as ‘Not do contracted’ in Figure 6, of which (40) is an example:

(39) Miss isn’t it that inflection do not change the meaning (S1B-015)

(40) Cos this guy don’t look like a true geek... (S1A-063)

Overall non-agreement rates with the verb do are highest in direct conversation and lowest in broadcast news, with class lessons ranging in between. That is, the rates decrease with increasing formality. While ‘Do’ non-agreement is restricted to conversations, non-agreement with the full negative form is only found in texts of medium formality. Example

69 The examples given in (38) through (40) are counted as instances of non-agreement. The verb in each case is not marked for third-person singular and thus does not agree with its subject. The corresponding cases of grammatical agreement would have been she does, inflection does not change, and this guy doesn’t look, respectively.
(39) is the only case of ‘Not do’ non-agreement in my spoken data. Looking at the source text it seems likely that the non-agreement has in fact been caused by a lack of plural marking in the noun rather than a non-standard use of the verb. In the text the speaker is asked to clarify what he means by uttering (39) and he rephrases his question as follows:

(41) I’m asking if it is that inflections do not change the meaning

Here the same utterance is reproduced and the only difference between the two sentences is that the noun is now marked for plural. Although the noun in (39) may have been an intended plural, this instance is nevertheless counted as non-agreement, because the decision has been made to focus on the surface structure of the nouns in question in order to avoid arbitrary interpretative decisions (cf. section 4.2.3).

What Figure 6 also shows is that by far the most frequent type of non-agreement with do is the invariant contracted form don’t used with singular subjects. Not do contracted\textsuperscript{70} with third-person singular subjects occurs 16 times in direct conversations, 7 times in class lessons, and not at all in broadcast news. Thus this feature decreases sharply across the registers with increasing formality levels in total numbers, but not in the amount of non-agreement that this phenomenon accounts for in the respective text category, namely 12.8%, 13.5%, and 0% in conversations, class lessons and broadcast news, respectively (cf. Figure 4 above).

Although the invariant contracted don’t (non-agreement form) is more frequent with singular subjects in direct conversations than is the agreement form doesn’t, it is not a general simplification feature in Jamaican English, not even in informal unmonitored speech. All speakers also use the standard form, which makes number distinctions—doesn’t for third-person singular and don’t for third-person plural subjects. The partial loss of subject-verb agreement that Trudgill (1990) found for some British dialects cannot be postulated for Jamaican English. There is merely a tendency to extend the use of don’t to third-person singular contexts in informal speech. The decrease in numbers of non-agreement with increasing formality clearly shows that non-agreement with do and more specifically the use of invariant don’t is a non-standard feature, but not a general simplification tendency in Jamaican English.

\textsuperscript{70} What we call Not do contracted in Figure 6 is the invariant negative form don’t used with all persons and numbers. Trudgill (1990) describes this as don’t simplification. As it is not clear at this point whether invariant don’t is in fact used in Jamaican English to simplify inflection, the term is not used in order to avoid premature evaluation of the phenomenon.
A phenomenon which qualifies as a plausible explanation for the high frequency of contracted don’t with third-person singular subjects in low formality texts is the presence of the invariant duont in mesolectal Jamaican Creole (Patrick 2004: 417f., see also Deuber 2009a: 21), an invariant form used as a negator with all persons and numbers. Deuber reports that he don’t, while being rather frequent in her 40 conversations, is not attested at all in ICE-GB and thus provides further evidence that this feature cannot be explained by general characteristics of non-standard varieties.

In class lessons, where the influence of the creole sharply decreases, agreeing contracted verb forms are far more frequent than are contracted verb forms that do not agree with the subject, and in broadcast news contracted verb forms are totally absent. This supports the assumption that invariant don’t is a salient overt creolism and thus tends to be avoided in formal speech. In all text categories in which the full positive do and negative (not) do occur, agreement is more frequent than is non-agreement.

Turning now to the distribution of non-agreement with different verbs, Figure 7 (a – c) shows the most frequent verbs following third person singular subjects in each category.
Figure 7 illustrates our previous finding that the rates of non-agreement decrease considerably with increasing formality, that is, that formality strongly influences the amount of non-agreement. While there is a large amount of non-agreement in low formality conversations, non-agreement is very infrequent in the category of highest formality, broadcast news, with class lessons ranging in between. The fact that there are different rates of non-agreement with the exact same verbs in conversations and class lessons, for example 29.2% non-agreement with the verb *come* in conversations (that is 7 tokens) versus 11.8% in class lessons (that is 2 tokens), proves that this is a stylistic difference, as speakers choose between standard and non-standard forms, depending on the perceived appropriateness of the respective forms under different circumstances, e.g. in formal versus informal situations.
What also becomes evident is that the single most frequent verb in all categories is *have*. This extremely high frequency is due to the fact that the overall numbers of the verb are analysed here irrespective of whether *have* is used as a main verb, as semi-auxiliary (*have to*) or as auxiliary verb. As expected, the rate of non-agreement with this verb in all its uses is highest in direct conversations (23.3%). In class lessons and broadcast news non-agreement with *have* is very infrequent, with rates of 3.3% and 6.6%, respectively. Figure 7 (a – c) shows the agreement patterns for *have* and the next five most frequent verbs in each category. A comparison of the figures provides further evidence of our finding that register influences the choice of vocabulary. Not only are collective nouns unevenly distributed in the three spoken registers (see also our discussion above), but the same is also true with regard to the frequency of particular verbs. Most obvious is the striking difference between conversations and class lessons on the one hand and broadcast news on the other. In conversations and class lessons four of the five most frequent verbs are the same, namely *don’t*, *say*, *come* and *happen*. Clear differences emerge, however, when more verbs are considered. In direct conversations, among the ten most frequent verbs we find several verbs of movement and recounting that are frequently used in narrative discourse, such as *happen* or *say*, *come* and *go*, *get*, and verbs of experiencing, such as *do*, *look*, or *sound*. This reflects the most prominent topics in this text category, topics which often contain narrative sequences, recounts of personal conversations, and gossip sessions, for example, or, in the case of the more interview-like texts in this text category, conversations about university, extra-curricular activities, or working life. In class lessons, many highly frequent verbs are related in some way to knowledge, such as *mean*, *ask*, or *equal*, the latter occurring very frequently in some classes in physics. *Become* and *happen* usually refer to historical developments.

By contrast, there is considerably more diversity in the use of verbs in broadcast news. Here verbs of reporting are most frequent, such as *say*, *report*, *state*, or *note*. Most of these occur in relatively fixed expressions such as *the police (who are investigating) say/report, the police continue investigations into, or RJR News understands that*. While the verbs in low formality texts are mostly simple, high frequency verbs, those in news are more formal, for example *obtain* versus *get*, or *state* and *report* versus *say*. In conclusion, it is

---

71 A closer analysis of whether the verb shows different agreement patterns in various kinds of uses is beyond the scope of the present study.

72 The label *don’t* includes all contracted negative forms of do, i.e. the plural form *don’t* and the singular form *doesn’t*. 
not only the overall frequency of non-agreement but also the frequency of particular verbs that is strongly influenced by register and the level of formality and that formality strongly influences non-agreement rates.

Statistical tests\textsuperscript{73} show that, in my spoken data, there is indeed a correlation between formality and overall rates of non-agreement with lexical verbs. The Chi square test ($\chi^2 = 81.1$) is highly significant (cf. footnote 59 above), and a logistic regression with grammatical non-agreement as the dependent variable and formality as the independent variable yields the results presented in Table 6.

<table>
<thead>
<tr>
<th>Variables in the Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>Formality</td>
</tr>
<tr>
<td>Constant</td>
</tr>
</tbody>
</table>

a. Variable(s) entered on step 1: Formality.

Table 6: Logistic regression for grammatical non-agreement and formality (spoken)

The variable B ($B = -0.876$ in step 1) indicates that there is a negative correlation between grammatical non-agreement and the level of formality, that is, with increasing formality the amount of non-agreement decreases. Formality is thus a good predictor of non-agreement, and the highly significant (Sig.=0.000) Wald test confirms this.

The chi square test for subject number ($\chi^2 = 107.0$) is also highly significant across my whole spoken data set. This means that non-agreement rates are significantly different when the subject is singular from when it is plural. While the Chi square value suggests that there is a correlation between the grammatical number of the subject and the amount of non-agreement in my overall spoken data, this is not true if only the category of broadcast news is examined, as the difference in non-agreement rates between sentences with singular as opposed to plural subjects is negligible in the text category of highest formality. Although the subject number is overall a good predictor of agreement according to the Wald test, the insignificance of the influence of subject number on the amount of

\textsuperscript{73} Due to the fact that lexical verbs and the exceptional verb be show very different agreement patterns and seem to be affected in different ways by formality and the various grammatical factors, statistical tests were computed separately for both verb types in order to avoid obscuring underlying linguistic constraints. The statistical tests for be are discussed in section 4.3.2.
non-agreement in broadcast news indicates that there may be an interaction between subject number and another factor. After the above discussion of grammatical agreement and non-agreement with lexical verbs in my spoken data it seems that the overall pattern can be explained by the effects of formality and register characteristics on the agreement patterns.

Last but not least, I tested the dependence of the potential explanations for non-agreement, such as unmarking, notional considerations or the proximity of a noun other than the controller of agreement, on formality and subject number. The cross-tabulation of potential explanation versus formality yields a Chi square approximation of $\chi^2 = 54.2$. This points to the probability of a correlation between formality and types of non-agreement. The contingency coefficient (0.487) shows that the correlation, as described above, is moderately strong.

The same test of dependence of the potential explanations of non-agreement on subject number yields an approximation of $\chi^2 = 15.464$ (df = 4), a result which is also significant at the 0.005 level, but the contingency coefficient (0.286) is much lower and points to a rather weak correlation between the type of non-agreement and subject number. This provides support for the assumption that there is indeed an interaction between two variables, that is, that one factor, presumably formality, causes the characteristic pattern of agreement in verbs following singular as opposed to plural nouns.

In the following section, subject-verb agreement with be will be examined across the three text categories representing low, medium and high levels of formality. Again, the types of non-agreement and the dependence of non-agreement on features such as subject number and formality will be the focus of attention, but additional potential influences, such as the tense of the verb, will also be analysed.

---

74 As frequencies must not be lower than 5 in any of the categories, there were too few instances for a Chi square test to be carried out. For this reason, SPSS computes a Chi square approximation. The Chi square result $\chi^2 = 54.2$ exceeds the critical value of 21.955 at the 0.005 probability level with df = 8. However, this Chi square result alone is not a reliable indicator of the presence or absence of a correlation, which is why the contingency coefficient must also be considered.

75 If the contingency coefficient $C > 0.2$ then there is a correlation, and the higher the value of $C$, the stronger is that correlation. A value of $C = 0.487$ points to a moderately strong correlation between the variables under consideration (cf. Walsh 1990: 191).
4.3.2 Agreement with the Verb BE

Table 7 shows the distribution of the verb *be* across the three levels of formality in the spoken and written data sets.

<table>
<thead>
<tr>
<th>Mode of Production</th>
<th>Spoken</th>
<th>Written</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formality</td>
<td>Tokens</td>
<td>%</td>
<td>Tokens</td>
</tr>
<tr>
<td>Low</td>
<td>1519</td>
<td>67.1</td>
<td>744</td>
</tr>
<tr>
<td>Medium</td>
<td>1667</td>
<td>54.5</td>
<td>1392</td>
</tr>
<tr>
<td>High</td>
<td>1147</td>
<td>53.0</td>
<td>1019</td>
</tr>
<tr>
<td>Total</td>
<td>4333</td>
<td>57.9</td>
<td>3155</td>
</tr>
</tbody>
</table>

Table 7: Distribution of the verb *be* across the spoken and written text categories

Of the 7488 occurrences of *be* in the whole data set 57.9%, that is 4333 tokens, were found in the spoken text categories direct conversations, class lessons and broadcast news, representing low, medium and high formality levels respectively, and 42.1% of all instances of *be*, that is 3155 tokens, were found in the written text categories of roughly corresponding formality, social letters, student essays and press news reports. There are always more tokens in the spoken than in the written category within the same level of formality. Of all the instances of *be* on the level of low formality, 67.1% are found in the spoken and 32.9% in the written data set. This comparatively low number of *be* verbs in written low formality texts, i.e. social letters, may be due to the smaller number of texts in this category, where there are only 15 texts as opposed to 20 in all other, spoken and written, text categories. On the level of medium formality, 54.5% of all instances of *be* occur in spoken versus 45.5% in written texts, and on the level of high formality, 53.0% were found in the spoken and 47.0% in the written texts. Only the spoken data will be considered in this section. Table 8 shows the distribution of agreement and non-agreement with *be* across the three spoken registers.

<table>
<thead>
<tr>
<th>Grammatical Agreement</th>
<th>Agree</th>
<th>Non-agr</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tokens</td>
<td>%</td>
<td>Tokens</td>
</tr>
<tr>
<td>Formality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>1507</td>
<td>99.2</td>
<td>12</td>
</tr>
<tr>
<td>Medium</td>
<td>1652</td>
<td>99.1</td>
<td>15</td>
</tr>
<tr>
<td>High</td>
<td>1129</td>
<td>98.4</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td>4288</td>
<td>99.0</td>
<td>45</td>
</tr>
</tbody>
</table>

Table 8: Distribution of agreement and non-agreement with *be* in the spoken registers
The majority of all sentences considered has singular subjects, only 20.4% of the instances of *be* in the category of lowest formality have plural subjects, 20.0% in medium formality and 31.7% in high formality texts. In comparison to the many hundreds of instances of grammatical agreement, non-agreement with *be* is extremely rare in the data, regardless of the number of the subject or the text category and register. Overall, only 1.0% of the instances of *be* do not agree with their subjects, which makes the verb *be* strikingly different from lexical verbs in this respect, for which a non-agreement rate of 9.8% was found (see section 4.3.1). While the overall non-agreement rates with *be* are negligible, the pattern can nevertheless reveal interesting facts about subject-verb agreement with *be* in educated Jamaican English.

Non-agreement is almost equally infrequent in low and medium formality texts with 0.8 and 0.9% respectively, and it rises slightly to 1.6% in high formality texts. A Chi square test ($\chi^2 = 4.369$, df=2, p=0.1169) reveals that there is no significant influence of formality on the rates of non-agreement with the verb *be* in my spoken data. This is strikingly different from the result for lexical verbs where formality was a highly significant factor influencing the amount of non-agreement.

As regards the influence of the grammatical number of the subject on the amount of non-agreement, the verb *be* again turns out to be different from lexical verbs, as non-agreement is higher with plural subjects in all categories. Non-agreement with singular subjects rises from 0.1% in conversations over 0.5% in class lessons to 1.0% in broadcast news. By contrast, non-agreement with plural subjects is found in 3.5% of all instances of *be* in conversations and 2.7% in both class lessons and broadcast news. A Chi square test for subject number across all spoken text categories ($\chi^2 = 47.986$ df=1, p=0.000) shows that subject number is a highly significant influencing factor on non-agreement rates with the verb *be*. This is not true for lexical verbs, as the discussion above has shown. This higher frequency of non-agreement with plural subjects can be partly accounted for by indirect influence from the mesolect, where invariant *iz* or *woz* is used as copula with all persons and numbers (cf. Patrick 2007: 139). This copula is marked for tense but does not have person or number distinctions. Its presence in the mesolect indirectly influences agreement with *be* in Jamaican English by facilitating non-standard use of *is* or *was* with plural
subjects especially in informal registers. However, not all non-agreement with *be* can be accounted for by the invariant mesolectal copula. Other factors that may be responsible for the higher frequency of non-agreement with plural subjects will be explored below.

Turning to the influence of tense on the non-agreement rates with *be* we find that non-agreement is slightly more frequent in present tense than in past tense when absolute frequencies are considered. This supports Deuber’s (2009a) finding that non-agreement is largely absent from past tense sentences in direct conversations, and slightly more frequent in sentences with present tense reference. However, if we take into consideration the much higher overall number of occurrences of *be* in sentences with present as opposed to past tense, it is misleading to regard only absolute token numbers. And indeed, if the instances of non-agreement are considered in relation to the total number of present versus past tense *be*, there is more non-agreement in past tense than in present tense. In conversation, of all present tense instances of *be*, 0.6% do not agree with their subject, versus 1.5% of all sentences with past reference; in class lessons there is 0.6% non-agreement in present tense versus 2.9% in past tense; and in broadcast news the percentages are 1.3% for present tense non-agreement and 1.9% for past tense. This shows that tense does exert an influence on the non-agreement rates with *be*. A Chi square test ($\chi^2 = 11.111, \text{df}=1, p=0.000$) confirms that tense is highly significant in determining rates of non-agreement with *be* in my spoken data.

Despite the small amount of non-agreement, the two factors subject number and tense have proven to be particularly strong influencing factors on the amount of non-agreement with *be*. This verb is thus crucially different from the lexical verbs analysed above. However, it is not yet clear whether this is also true with regard to the types of non-agreement, i.e. notional agreement, proximity agreement, etc.

An examination of the types of non-agreement that occur may prove insightful if compared to the types that occur with verbs other than *be*. If the types of non-agreement are similarly distributed across the categories when the verb *be* is used rather than one of the other verbs analysed above, this distribution would point to general tendencies of non-agreement that do not depend on the type of verb but on other factors, such as subject number or formality and register. Figure 8 gives an overview of the types of non-agreement in each register; the percentages refer to percent of all non-agreement in that given text category:

---

76 The forms *I* is, *you* is and *they* was are attested in the informal spoken section of ICE-Jamaica, albeit only once. Their infrequency points to the high salience and stigmatization of these forms in Jamaican English.
A pattern emerges similar to that which appeared in the analysis of other types of verbs, namely that notional agreement is highest in texts of high formality and accounts for most of the non-agreement in this category. However, with direct conversations and class lessons the pattern is reversed. With the verb *be* notional agreement is more frequent in class lessons than it is in direct conversations, whereas this is not true for other verbs. This is because, in contrast to lexical verbs, with *be* there is not a single case of non-agreement with an indefinite pronoun subject in my conversational data, although such cases accounted for most of the instances of notional agreement with lexical verbs.

While the rate of notional agreement is around 18% in direct conversations, 40% of all non-agreement in class lessons and around 67% in broadcast news can be attributed to notional considerations, an example of which is given as (42).

(42) **Guyana are** Red Stripe Bowl champions for the second time (S2B-018)

This means that in these cases it is not the grammatical number of the subject that determines agreement but instead the noun is given a collective reading. A high proportion of the non-agreement involving notional considerations, particularly in highly formal texts, consists in examples such as (43) in which a sum of money is interpreted as the total sum rather than the individual dollars and thus takes a singular verb.

(43) **Five point two million US dollars** in foreign exchange **was** purchased by the island's commercial banks yesterday
The tendency in informal texts to extend *is* and *was* to plural contexts together with the frequency of this type of notional agreement in formal texts explains why there is more non-agreement with *be* with plural rather than singular subjects. Proximity agreement with *be* is totally absent from low formality texts but accounts for about 7% of all non-agreement in medium formality texts and for about 22% in high formality texts. An example is given in (44):

(44) **Preparations** for the **ceremony** *is* now in high gear (S2B-020)

Here, the verb agrees with the more proximal singular noun *ceremony*, although the syntactic head of the subject noun phrase is the plural noun *preparations*. The frequencies found in my data do not support the first part of Quirk’s observation that “[p]roximity concord occurs mainly in unplanned discourse. In writing it will be corrected to grammatical concord if noticed” (Quirk et al. 1985: 757). The analysis instead yields results similar to those of Sand and Deuber, which show that proximity agreement is extremely infrequent in informal, spoken direct conversations (Deuber 2009a) and more frequent in news texts (Sand 1999). This is true for both *be* and other types of verbs.

While the amount of proximity agreement correlates positively with the level of formality, it is, however, not so much the overall frequency that enables us to draw conclusions about the potential influence of formality on this type of non-agreement. In other words, this correlation cannot to be taken as evidence of a direct dependence of proximity agreement on register. The correlation might be caused by other factors, such as the number of complex noun phrases, which tends to be higher in more formal texts due to the fact that these texts usually have a higher lexical and informational density than do informal conversations. For this reason, the frequency of proximity agreement always has to be considered in relation to the frequency of complex noun phrases consisting of a head noun and a local noun, i.e. those noun phrases where proximity agreement would in principle be possible. An analysis of all the sentences with lexical verbs in my spoken data shows that in the text category direct conversations the number of such complex noun phrases is rather low, only 27, while the number increases with increasing formality to 53 in class lessons and 97 in broadcast news. The complexity of noun phrases thus increases with formality. Proximity agreement is found in 7.4% of all complex noun phrases in direct conversations (2 tokens), in 3.8% in class lessons (2 tokens), and 3.1% in broadcast news (3 tokens, including the one instance of hypercorrect marking which can be attributed to the proximity of a singular local noun). The analysis suggests that proximity agreement
increases with increasing formality when total numbers are concerned but the pattern is reversed if the number of complex noun phrases is taken into consideration.

As the results so far were obtained from analyses of spoken data only, it is not yet possible to make assertions about the validity of the second part of Quirk’s assumption or about the validity of Sand’s observation that proximity agreement occurs equally frequently in spoken and written language. These hypotheses will be discussed in detail below, in section 4.5.

Verb number refers to instances of non-agreement where the verb does not agree with the subject noun in number, but no alternative explanations, such as proximity of another noun or notional number, can be found. In the overwhelming majority of cases, the subject number is plural and the verb used is singular, as in (45), but occasionally a plural verb is used with a singular subject (46).

(45) Her parents was upset was upset... (S1A-061)

(46) Protocol are a treaty or a convention (S1B-008)

This type of grammatical non-agreement decreases from about 54% in direct conversations to 33% in class lessons and down to 6% in broadcast news. Much of this variation can be explained by the invariant copula iz or woz mentioned above, which functions as the copular form for all persons and numbers in mesolectal Jamaican Creole (Patrick 2007: 139) and is here extended to plural contexts in ways that are not possible in International Standard English, or by notional considerations (as described in section 4.3.1). The decrease might then partly be due to the lower influence of Jamaican Creole in texts of higher formality.

A closer analysis of subject-verb agreement with be in a much larger sample than the present one is needed to gain more reliable insights into agreement with be. For example, the data do not allow an investigation of the influence of polarity on non-agreement with be, because there are only four sentences of negative polarity, three of which occur in existential there + be constructions (ETBs), the last phenomenon to be briefly mentioned here.

This type of non-agreement accounts for about 18 per cent of non-agreement with be in direct conversations, 20% in class lessons, and 5.6% in broadcast news. The numbers are

77 The category “other” consists of verbs occurring in phrases that display syntactic structures like coordinated subjects, which may influence agreement. As such ‘other’ cases are very infrequent they will not be discussed in detail.

78 Repetitions are excluded and this example counts as one instance of non-agreement only.
rather low, but an analysis of only existential constructions reveals some interesting patterns. First, as Figure 9 indicates, the non-agreement rate is consistently lower than 10% in all three text categories, but it decreases steadily with increasing formality.

![Figure 9: Agreement and non-agreement in existential there + be constructions](image)

In my conversational data, 4.85% of all existential *there + be* constructions are cases of non-agreement. The rate then decreases to 1.5% in class lessons and 0.9% in broadcast news. However, the numbers are too low for statistical significance tests, and the apparent tendencies reported here will have to be tested in future research on the basis of a larger sample.

While it seems that formality is the crucial influencing factor here, there need not be a causal relation between both variables. The figure merely indicates a slight negative correlation between non-agreement and the level of formality, that is, the rate of non-agreement decreases with increasing formality.

Alternative explanations for this agreement pattern might be found by considering characteristics of the registers that are here associated with a particular (relative) level of formality, factors such as the higher syntactic complexity and lexical density of utterances in class lessons and broadcast news as compared to that in direct conversations or the higher cognitive load in lectures, which Crawford (2005) found to be highly influential on agreement in existential constructions. The author compares ETBs (i.e. existential *there + be* constructions) in spoken and written texts of different levels of formality in order to  

---

79 Crawford (2005) investigates only present tense *there + be* constructions and does not discuss potential influences of tense on (non-) agreement.
provide a thorough analysis of this construction across a wide range of text types, including in his considerations also factors such as public or private, size of audience, and the relationship of the speaker/writer to the audience. The text categories covered are face-to-face conversations and academic lectures as spoken registers of different relative levels of formality, and a corpus of academic textbooks, a fiction corpus and a corpus of chat room language as written registers. In contrast to previous assumptions, Crawford argues that agreement mismatch in *there + be* constructions is not due to informal, conversational language use (cf. Crawford 2005: 39) but that it is in fact more frequent in more formal spoken registers. His data reveal, first, that non-agreement in ETBs is indeed more frequent in spoken than in written language use (p. 42); second, that the subject number is important, as “lack of concord is nearly always found with a singular form of *be* and a plural notional subject” (p. 42); third, that non-agreement is most frequent with the contracted form of the singular verb, *there’s*, preceding a plural complement, and less frequent with the full verb, *there is* (p. 45); and fourth, that there is a higher percentage of non-agreement with the contracted form in lectures than in face-to-face conversations (p. 46). He attributes this, on the one hand, to the higher frequency of plural notional subjects in lectures than conversations, and, on the other hand, to the higher cognitive load in lectures (Crawford 2005: 49), where existential *there + be* constructions are used more frequently than in conversations for discourse management functions. Crawford concludes that the spoken versus written dichotomy is a better indicator for non-agreement in *there + be* constructions than formality.

In my spoken data, the overall frequency of existential *there + be* is highest in class lessons (111 tokens) and lowest in broadcast news (33 tokens), with direct conversations ranging in between the two (77 tokens). These results provide evidence for Crawford’s assumption that existential *there + be* constructions occur more frequently in lectures than they do in direct conversations. Following Crawford, the cognitive load can be assumed to be highest in lectures80 and thus might cause the high number of existential constructions in the category of class lessons. This is because, on the one hand, lectures are not pre-formulated in the way news texts are—news texts are planned and edited before the actual speech event begins. The utterances in lectures have to be produced on the spot and the speaker has to keep track both of what has been said before and where the line of argumentation is going, or, in other words, what arguments s/he should choose next and

---

80 See Crawford (1999) for a more detailed discussion on the cognitive load in lectures as compared to that in direct conversations. In this paper, only a very short description of influencing factors can be given.
how s/he should present them. On the other hand, lectures are not subject to the competitive strategies of turn keeping and turn taking as are unplanned casual conversations, thus the lecturer has more time to plan the next utterance and to formulate syntactically complex and lexically dense utterances. The necessity of managing discourse in this way may cause existential *there + be* constructions to be used rather frequently for summarizing topics or introducing new issues.

In contrast to the overall number of occurrences, however, non-agreement in existential *there + be* constructions does not seem to be affected by such differences in function; the percentage of non-agreement decreases steadily with increasing formality of the text, from direct conversations to class lessons to broadcast news (cf. Figure 9). My data show that non-agreement rates are not particularly frequent in the category of class lessons, possibly owing to the fact that this text category comprises not only lectures but also more interactive class lessons, such as discussions or tutorial groups.

Figure 10 shows the amount of non-agreement in existential *there + be* constructions in my spoken data depending on whether the full or contracted form of *be* is used.

![Figure 10: Non-agreement in ETBs by contractedness](image)

An example of non-agreement with the full form is given as (47) and with the contracted form as (48).

(47) You’re not allowed to put a lot of things in the fridge because *there is so many persons* living on hall (S1A-046)

(48) It sounds so far that *there's not really a lot of roles* for black people (S1A-062)
Indeed there seems to be an effect of contraction on non-agreement in educated Jamaican English, in that non-agreement is higher in contracted *there + be* constructions (7.7%) than in non-contracted ones (1.4%). My data thus support Meechan and Foley’s (1994) findings that contractedness is a factor exerting influence on subject-verb agreement in existential *there + be* constructions. Again, a systematic analysis of more homogeneous text types and a larger data base would be valuable and might yield important insights into the highly complex pattern of subject-verb agreement in existential *there + be* constructions.

In sum, the verb *be* behaves rather differently from regular verbs, as can be seen in the fact that, first, rates of non-agreement are so low as to be insignificant and second, non-agreement is more frequent with plural subjects than it is with singular subjects. While subject number is a highly significant factor influencing non-agreement with *be*, formality is not. This is strikingly different from the results obtained for lexical verbs, where formality has been identified as the better predictor of non-agreement rather than subject number (see our discussion in section 4.3.1 above). The difference in pattern can be explained by the different nature of non-agreement with the respective verbs. Subject number and formality interact in determining the amount of non-agreement with lexical verbs, as formality affects the presence or absence of inflectional marking, which generally occurs only on verbs that agree with singular subjects. This, however, is not the case for the verb *be*, as number is not inflectionally but morphologically marked.

The analysis of the types of non-agreement, however, yields a pattern similar to that found for lexical verbs, namely that the frequencies of notional and proximity agreement increase with increasing formality. Simple non-agreement of verb number and agreement mismatch in existential *there + be* constructions, on the other hand, decreases with increasing formality. *There + be* constructions have been found to occur more frequently in class lessons than in other text types, but the effects of register on non-agreement in such existential constructions across the categories were not as strong as those described by Crawford (2005), a difference which may be due to the heterogeneity of my categories.

### 4.3.3 Conclusion – Subject-Verb Agreement in Spoken Jamaican English

In my analysis of subject-verb agreement in three spoken text categories of ICE-Jamaica, direct conversations, class lessons and broadcast news, a substantial amount of variation has emerged with regard to phenomena that have frequently been identified as
characteristic features of Jamaican English. Subject number is a factor that exerts a strong influence on (non-) agreement with the exceptional verb be, regardless of register, as well as with lexical verbs, in which case, however, it has been found to interact with the factor formality. As described above, the type of verb, i.e. whether it is a lexical verb or the verb be, also plays a role in determining agreement. In addition, tense has been identified as another crucial factor influencing agreement with be.\(^{81}\)

However, the more interesting and complex findings appear in a systematic analysis of the prevailing types of non-agreement in the categories under consideration. In my data, evidence was found for Deuber’s (2009a) assumption that a lack of inflection on verbs following third-person singular subjects is a frequent phenomenon in informal Jamaican English. Moreover, a slight tendency to extend notional concord appeared in my informal data supporting Sand’s (1999) results. However, these features did not prove to be evenly distributed in all registers; there was no unidirectional tendency for all types of agreement to decrease with increasing formality. In fact, four trends appeared in my data which contribute to the rather complex pattern of subject-verb agreement in Jamaican English:

First, overall rates of grammatical non-agreement decrease with increasing formality; grammatical non-agreement is most frequent in direct conversations and least frequent in broadcast news. This finding, together with the occasional hypercorrectly marked verbs in the higher formality levels, indicates that leaving verbs unmarked is stigmatized, while the use of -s marking on third-person singular verbs indicates speaker prestige.

Second, register correlates differently with the various types of grammatical non-agreement. The absolute number of proximity agreement, for instance, increases with increasing formality and thus definitely cannot be characterized as a feature of unplanned discourse (cf. Quirk et al. 1985). However, as the information density is higher in formal than informal texts, the complexity has to be taken into account. When proximity agreement is analysed with regard to the number of sentences containing complex nouns, Quirk’s assumption has been shown to hold true for the spoken data, as the percentage of proximity agreement in sentences with complex noun phrases decreases with increasing formality.

The amount of notional agreement also tends to increase with increasing formality, but here the pattern is not straightforward. While the tendency to extend notional agreement to indefinite pronouns such as somebody or everybody was found in direct conversations (see \(^{81}\) In the case of lexical verbs, tense was not considered as number agreement is marked only on verbs in the present tense.
also Deuber 2009a), it was rather infrequent in class lessons and completely absent from broadcast news. The finding that the rate of notional agreement is nevertheless higher in news texts can be explained by the higher number of collective nouns in this category. Here, register characteristics are extremely influential in that they account for an opposing trend, particularly regarding the choice of vocabulary. While the high frequency of indefinite pronouns in direct conversations and their absence in news texts can be explained by the prevalence of interpersonal topics in conversations and the lack of such topics in news texts, the high number of collective nouns in news also can be attributed to common news topics. In broadcast news, topics such as politics, where the nouns government and ministry are particularly well represented, or economics, where recurrent reference is made to company, are much more frequent than in conversations, where collective nouns, such as family, do occur, but much more rarely. This register dependence was also reflected in an analysis of the most frequent verbs in each category. Narrative verbs and verbs of experiencing occurred much more often in direct conversations than in the other text types, while academic, knowledge-related verbs were most frequent in class lessons, and reporting verbs were most common in news.

Third, a simplification process of the type Trudgill (1990) found in dialects of England, or the general trend in varieties of English towards erosion of inflectional morphology, does not qualify as a plausible explanation for agreement variation in my data, as these features do not consistently occur in all categories. What Trudgill calls don’t simplification, that is, the use of invariant negative contracted don’t regardless of the number of the subject, occurs mostly in direct conversations and occasionally, but not very frequently, in class lessons. However, such instances are not found in broadcast news texts. Thus, one cannot speak of a general trend in Jamaican English to simplify inflection. The more likely explanation for the extension of the contracted form don’t to third-person singular subjects can be found in the influence of the creole form duont, which can be used with all persons and numbers. Creole influence is the most probable explanation for most of the types of non-agreement in the category of direct conversations, namely the extension of notional agreement to indirect pronouns, the use of the quantifier much with plural count nouns, the use of is, was and don’t with plural subjects, and last but not least, the simple unmarking of verbs.

However, and this is the fourth trend revealed by my data, creole influence decreases strongly with increasing formality. In direct conversations, lack of inflection occurs particularly frequently in environments where other creole features can be found, such as
the use of him as subject pronoun. Both the number of such creole environments and the instances of unmarked verbs decrease in class lessons and decrease even more in broadcast news, where creole environments are no longer found and unmarked verbs are rather scarce. Thus, it appears that Mair’s (2002) conclusion that “Patois is clearly the dominant shaping influence on spoken English in Jamaica” can be fully supported in the category of direct conversations, is less supported in the category of class lessons, and does not seem to hold in the case of broadcast news, since in this last category notional agreement is far more influential than are creole features. The analysis indicates that conversations and class lessons resemble one another in many respects when they are compared to broadcast news. This is true for the prevailing types of non-agreement and for occurrences of different kinds of notional agreement, but also for the most frequent verbs and the rates of don’t simplification and proximity agreement. The contrast between conversations and class lessons is much weaker than that between class lessons and broadcast news. The heterogeneity of the categories that were the basis for this study might account for this similarity, as the text type, direct conversations, comprises not only casual conversations but also more formal interviews, and class lessons include not only rather formal lectures, but also more interactive and informal tutorial groups and class discussions. The texts within the category broadcast news, however, are more homogeneous and consistently on a highly formal level. Moreover, broadcast news consists of scripted texts, while all others are produced in time. Thus this text type contrasts sharply with the other two, while the differences between the other two text types are not so marked.

In order to be able to compare spoken to written language use and to draw conclusions as to how much variation in agreement patterns can be attributed to the mode of production, the next section presents an analysis of subject-verb agreement in three written text categories representing low, medium and high relative levels of formality, namely social letters, student essays, and press news reports. The results from the spoken data set are left aside here, and both the findings from spoken as well as written data are compared below in section 4.5. The analysis of the written data will concentrate on the same aspects as that of the spoken data presented above, namely lexical verbs and the exceptional verb be, the effects on the amount of non-agreement of formality, subject number and the types of non-agreement as well as their potential explanations. In the case of be, tense will be considered as an additional, potentially influential, factor.
4.4 Results for Written Jamaican English

4.4.1 Agreement with Lexical Verbs

As mentioned above, 2033 lexical verbs were analysed in the written section of ICE-Jamaica. The 376 verbs in the informal category ‘social letters’ represent 18.5% of all written lexical verbs, the 1030 verbs in ‘student essays’ add up to 50.1% and the 627 verbs in the highly formal category ‘broadcast news’ constitute 30.8% of verbs other than *be*. The lexical verbs are thus distributed rather unevenly across the three written registers, on the one hand owing to the fact that the category of social letters comprises only 15 texts, whereas all other text categories consist of 20 texts, and on the other hand to the fact that the texts in ‘broadcast news’ have a higher information density and therefore exhibit more and more complex noun phrases.

Table 9 gives the token numbers and percentages of grammatical agreement and non-agreement in the spoken categories of low, medium and high formality respectively.

<table>
<thead>
<tr>
<th>Formality</th>
<th>Agree Tokens</th>
<th>Agree %</th>
<th>Non-agr Tokens</th>
<th>Non-agr %</th>
<th>Total Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>336</td>
<td>89,4%</td>
<td>40</td>
<td>10,6%</td>
<td>376</td>
</tr>
<tr>
<td>Medium</td>
<td>942</td>
<td>91,5%</td>
<td>88</td>
<td>8,5%</td>
<td>1030</td>
</tr>
<tr>
<td>High</td>
<td>610</td>
<td>97,3%</td>
<td>17</td>
<td>2,7%</td>
<td>627</td>
</tr>
<tr>
<td>Total</td>
<td>1888</td>
<td>92,9%</td>
<td>145</td>
<td>7,1%</td>
<td>2033</td>
</tr>
</tbody>
</table>

Table 9: Distribution of agreement and non-agreement with lexical verbs in the written registers

The overall rate in the 55 written texts is 7.1% for non-agreement and 92.9% for agreement. Non-agreement is thus quite infrequent in written Jamaican English. The rate of non-agreement is highest in texts of lowest formality, as 10.6% of all verbs in social letters do not agree with their subjects, and it decreases to 8.5% in medium formality texts and to only 2.7% in press news reports. A Chi square test of significance ($\chi^2 = 28.577$, df = 2, p= 0.000) proves that the influence of formality on non-agreement is highly significant.

In Figure 11, the percentages of grammatical non-agreement are shown separately for verbs following singular subjects and plural subjects across the three levels of formality.
In social letters, non-agreement rates are higher when the subject is singular, but the same is not true in student essays and press news reports. While the rate of non-agreement in verbs following singular subjects decreases steadily with increasing formality, from 12.1% in social letters to 8.4% in essays down to 2.4% in press texts, the rate of non-agreement in verbs following plural subjects is rather low in informal letters, with 5.7%, it then increases to 8.8% in medium formality, and decreases again to 3.1% in highly formal press texts.

Explanations for this pattern can only be found once the types of non-agreement are taken into consideration, but a Chi square test across the whole written data set ($\chi^2 = 0.679$, df=1, p=0.409) shows that subject number does not influence the overall rate of non-agreement to any significant extent. Formality is a much stronger factor influencing non-agreement rates than is the grammatical number of the subject.

The types of non-agreement prevailing in the three written text categories are given in Figure 12. In accordance with Figure 4 above, ‘Absent –s’ denotes the absence of singular inflection in verbs, ‘Do negative’ represents the invariant use of contracted negative don’t in the present tense irrespective of subject number, ‘NP-marking’ means that there is non-agreement within the noun phrase due to the head noun lacking plural marking, and ‘Hypercorrect’ denotes verbs following plural subjects which nevertheless have singular –s marking. The category ‘Other’ includes all those cases which cannot be classified under any of the types mentioned above.
Figure 12 indicates that non-agreement of the ‘Other’ type in each category is negligible if compared to the amount of non-agreement of other types, such as hypercorrect inflection and the absence of –s marking. It results in each category from coordinated subjects being followed by a singular rather than plural verb, as the example from social letters (49) illustrates.

(49) …the demands of this language, my masters, numerous part time jobs and my daughter's arrival in April, leaves me very little free time (W1B-005)

The most common type of non-agreement in written Jamaican English is the absence of singular inflection in verbs following singular subjects. This accounts for 72.5% of all non-agreement in social letters (29 tokens), 54.6% in student essays (48 tokens) and 52.9% in press news reports (9 tokens). This result provides further evidence for the assumption that the dependence on formality of the rate of non-agreement involving an absence of –s marking and the tendency of such non-agreement to decrease with increasing formality is a stable, general trend (cf. the conclusion concerning the spoken data) which is not influenced by the mode of production.

Non-agreement involving the negative contracted don’t does not decrease sharply with increasing formality, as it does in the spoken data, but it is restricted exclusively to social letters, the most informal of the written registers under consideration. Once, this type of non-agreement occurs in a cluster of creole features (50).
(50) …but i don’t know what wrong they keep turning him down… saying him don’t have enough "EXPERIENCE"… but you know if no one employs him how him going to get experience?\textsuperscript{82} (W1B-015)

In this example, several features can be found that have been identified as characteristic of the mesolectal range of the Jamaican creole continuum (see Sand 1999 and Deuber 2009a, among others), for instance the absence of the copula in what wrong and how him going to get…, and the use of him as a subject pronoun (cf. section 4.3.1).

The occasional appearance of non-agreement of negative contracted do in the data is most likely due to the invariant mesolectal form duont (see Patrick 2004: 417 and Deuber 2009a: 21). A brief corpus search of he don’t, she don’t and him don’t in several varieties of English\textsuperscript{83} confirms that it is a characteristic feature of Jamaican English rather than a general simplification of the type found by Trudgill (1990) in some British dialects or a tendency in post-colonial Englishes, i.e. an angloversal. The fact that invariant don’t is restricted to informal texts suggests that it is to be regarded as a direct and highly salient creolism which, despite being stigmatized, is sometimes used in English as a stylistic alternative.

Its occurrence in the written data shows that the influence of Jamaican Creole is not restricted exclusively to spoken language. Such direct contact effects are, however, rather infrequent in my written data and thus the creole cannot be assumed to affect written Jamaican English to any considerable extent. As expected, such invariant contracted forms are found neither in the highly formal press news reports nor in the student essays, i.e. texts of medium formality, providing even further evidence for its creole associations.

Non-agreement involving a noun phrase internal mismatch of number occurs twice in texts of low formality, five times in texts of medium formality and once in highly formal texts. In roughly half of these instances, the noun lacks plural marking, an example of which is given as (51). In the other cases the noun is marked for plural (52), although the phrase as a whole is clearly not intended as plural.

(51) Anyway the student have started to stream in so later. (W1B-004)

\textsuperscript{82} Some of the annotations have been removed for clarity, i.e. only the original text version is presented here.

\textsuperscript{83} The search was restricted to the spoken informal sections of the ICE-corpora discussed in chapters 5 and 6, as it is in informal spoken language where language change, e.g. the tendency toward inflectional simplification, tends to set in first. The total frequencies for he don’t, she don’t and him don’t are as follows. BritE: no occurrences, AmE: 1, IndE: 3, SingE: 7 and JamE: 28.
The Maroons (slaves who fled to the mountains after the British ousted the Spaniards; the names derives from the word <foreign>Cimarron</foreign>…

In each of the three written text categories, noun phrase non-agreement accounts for five to six percent of the non-agreement. As there are no clear differences here in the three levels of formality, this type of non-agreement will not be analysed in any more detail.

The last type of non-agreement to be discussed is hypercorrect marking of verbs following plural subjects. In contrast to the results in the spoken text categories, where hypercorrect marking was very infrequent both in texts of low and medium formality and very frequent in highly formal texts, the written data reveal a different pattern. Here, non-agreement involving hypercorrect marking is also infrequent in informal social letters, the two instances accounting for 5% of all non-agreement in this category, and comparatively high in the highly formal category of press news reports, in which the 6 tokens add up to 35.3% of all non-agreement. So far, the pattern is still in accordance with the results from the spoken data. However, the rate of hypercorrect non-agreement in the text category of medium formality is much higher in the written than in the spoken data. Whereas the rate is 3.8% in class lessons, hypercorrection accounts for 38.6% of all non-agreement in student essays.

Now the explanation of the above-mentioned unexpected agreement pattern with relation to subject-number becomes clear. The surprisingly strong increase of non-agreement following plural subjects in student essays corresponds to an equally strong increase in the number of hypercorrectly marked verbs in the same text category. The stark difference in non-agreement rates between spoken and written texts of medium formality can be plausibly explained by the tendency for written language to follow the educational target variety more closely than does spoken language, as the production of written text can be carefully planned beforehand. In other words, students learn that –s marking on verbs is a prestigious feature of Standard English, the variety aimed at in educational contexts (cf. also the discussion above in section 4.3.1) and hence they produce many hypercorrect forms in their university assignments and examination essays, but not in private letters.

The fact that language in class lessons is spoken and produced on the spot without much time for planning accounts for the fact that only few hypercorrect forms occur in this

84 The effects of educational targets on patterns of hypercorrect non-agreement in university contexts supports the view that Jamaican English is influenced by effects of Second-Language Acquisition.
educational text type. A substantial number of hypercorrectly marked verbs can also be found in the formal category of press news reports, although these texts are normally edited before being published. This provides additional evidence for the prestige carried by these forms.

These results suggest that the factor subject number strongly interacts with formality, as the latter factor determines the pattern with relation to the first. Only in informal texts do sentences with singular as opposed to plural subjects differ considerably, but with higher formality this difference disappears, since the numbers of unmarked verbs decrease with increasing formality, while simultaneously the number of hypercorrectly marked plural verbs increases with increasing formality.

Three of the hypercorrect forms in student essays may have been marked for singular due to the notional interpretation of the noun phrase. An example is given as (53).

(53) If one should consider the Human Development indicators that was published by the United Nations that ranks a country's development in terms of social development of a population (W1A-007)

Here, the Human Development indicators is interpreted as denoting the published document, consisting of an entire list of indicators, rather than as the individual indicators themselves, therefore a singular verb form follows. None of the instances of hypercorrect marking in social letters or press news reports can be attributed to notional considerations.

Eleven of the hypercorrections in student essays can be explained by proximity agreement, e.g. (54), while only two in press news reports and none in social letters can be so explained.

(54) Changes in the way we communicate and do business has affected mankind both positively and negatively. (W1A-014)

Here, the head of the subject noun phrase is the plural noun changes, i.e. for grammatical agreement the verb would have to be plural, but the verb agrees with the more proximal noun business, and therefore displays singular –s marking. In written texts of higher formality, i.e. press news reports, only few instances of hypercorrection can be found. The fact that hypercorrect proximity agreement is edited out of the carefully prepared press texts but appears rather frequently in educational texts suggests that it is an indirect contact influence which is even less salient than zero forms.
In a next step, the type of non-agreement accounting for most of the non-agreement in my written data is taken into consideration, namely the absence of singular inflectional marking in verbs following singular subjects.

The strongest factors causing the absence of –s marking in verbs are shown in Figure 13.

![Figure 13: Potential explanations of non-agreement involving absence of –s marking across formality levels (written)](image)

In the written data, absent –s marking occurs only once in a context where other creole features are present, and that is in the category of lowest formality, social letters.

Four of the cases of non-agreement in student essays are caused by a lack of plural marking in the noun leading to non-agreement, as in (55), which is clearly intended as plural if the context is taken into consideration.

(55) ...that student who can not afford it, but have the ability will still be able to get a University education. (W1A-002)

In press news reports there is only one such instance of non-agreement caused by missing plural noun marking.

Proximity agreement is totally absent from social letters and accounts for 25% of all non-agreement involving absence of –s marking in student essays and 22.2% in press news reports. Thus Quirk et al.’s (1985) assumption that proximity agreement “will be corrected to grammatical concord if it is noticed” is not confirmed for the written data when the token numbers are concerned. However, the higher proportion of proximity agreement in texts of higher formality in both spoken and written language is not necessarily a direct
consequence of formality but it can be attributed to the higher lexical density that is a
typical characteristic of formal language, that is, the tendency of formal texts to contain
more complex noun phrases than informal texts. In fact, the number of complex noun
phrases in sentences containing lexical verbs is very low in social letters, where only 26
tokens were found, and it increases dramatically to 187 in student essays, the category of
medium formality. In press news reports some 138 complex noun phrases were found. In
the informal category of social letters there was not a single case of proximity agreement.
This is in stark contrast to student essays, where 12.3% of all complex noun phrases led to
proximity effects (23 tokens, including 11 instances of hypercorrect marking which can be
attributed to the proximity of a singular local noun). In press news reports, proximity
agreement was found in 2.9% of the clauses containing complex noun phrases (four
tokens, including 2 hypercorrectly marked verbs following singular local nouns). Student
essays are exceptional both with regard to the number of complex noun phrases and the
percentage of proximity agreement within these clauses containing complex noun phrases.
A reason might be that in the educational context of student essay writing the target is a
very formal writing style, and that students might ‘overdo it’ by using extremely high
numbers of complex noun phrases. Thus the register characteristics are very influential
here and it is not formality alone that exerts an influence on the agreement patterns
revealed by the data. About half of the instances of proximity agreement are hypercorrect
additions of singular inflection on verbs, and this high number can be explained by verbal
inflection being a form of high prestige (cf. the discussion above).

In press news reports, the percentage of proximity agreement in sentences with complex
noun phrases (2.9%) is higher than in informal written language but very low in
comparison to the texts of medium formality. It thus seems that Quirk’s assumption of
proximity agreement being “corrected to grammatical concord if noticed” holds true for
published texts in the context of the media, but not in the educational sector.

Figure 13 reveals a very clear pattern with regard to simple unmarking of verbs and
notional agreement. While simple unmarking of verbs decreases steadily from 79.3% in
social letters down to 43.75% in student essays and further down to 11.1% in press news
reports, the pattern for notional agreement is reversed. It increases from 17.2% in social
letters to 22.9% in student essays and up to 55.6% in press news reports.

A search of Hundt’s (2006) 35 collective nouns and the additional 7 nouns (see section
4.3.1 above) reveals that, in the written data, these are most frequent in class lessons (691
tokens) and least frequent in social letters (113 tokens) with press news reports ranging in
between (517 tokens). In contrast to the spoken data described above, not a single case of notional agreement in my written data is caused by non-agreement with indefinite pronouns such as *everybody*. This explains the relatively small amount of notional agreement in informal written as compared to informal spoken language and the more even increase in notional agreement with increasing formality in the written data.

What the analysis of different types of notional agreement in my written data also shows is that Deuber’s assumption of a tendency in Jamaican English to extend notional agreement cannot be supported for written language at all, as no cases of non-agreement caused by plural notional interpretation of indefinite pronouns are recorded, neither in informal social letters nor in texts of higher formality. In the few instances in the written texts where indefinite pronouns occur at all, they are always “unambivalently singular” (Quirk et al. 1985: 764). The same is true for the quantifier *much* which has been found in spoken language to occasionally determine plural count nouns in informal text types. In written Jamaican English, only a single instance of *much* + plural noun was found (56), and against any expectations, it occurred in the most formal text type of press news reports.

(56) Coleridge Barnett, principal of Wolmer’s Boy’s School, noted that *much of the free place winners* entering the school were in need of remedial classes,…

(W2C-004)

Based on the results described above one cannot speak of a tendency in written Jamaican English to extend notional agreement. Almost all instances of non-grammatical agreement occur with nouns (such as *government, management, university*, names of sports teams, etc.) that can be interpreted even in International Standard English as entities or groups of persons, hence singular in number, or as the individual members of the category, hence plural in number.

Looking at the most frequent collective nouns in each text category we notice again the effects of register characteristics. In social letters, the most frequently occurring collective nouns are *class, couple and family*, reflecting the private nature of the topics prevailing in this category. In the student essays, which are mainly concerned with university education, the single most frequent collective noun is *university*, which is found 366 times in the written category of medium formality, followed by *government* and *community*. The three most frequent collective nouns in press news reports are *government, company* and *ministry*, the same collective nouns found most frequently in spoken news texts. This result again points to the characteristic distribution of topics in the three written text categories.
and thus provides evidence that the strong influence of formality and register characteristics on the choice of vocabulary and syntactic structures postulated above for spoken language is true also in written language.

The verb *do* in present tense was found 98 times in my written data, in 92.9% (91 tokens) agreeing and in 7.1% (7 tokens) not agreeing with their subject. The majority of non-agreements occurred in sentences where the subject was singular, while only one case of non-agreement was recorded in a plural sentence (57).

(57) **A lot of the kids** in my class (not classes) **doesn’t** like me (W1B-003)

As non-agreement is most frequent with singular subjects, only these cases will be considered further below. Figure 14 shows the distribution of agreement and non-agreement with the verb *do* in positive versus negative and in full versus contracted forms for the three text categories representing low, medium and high levels of formality.

One important finding emerging from Figure 14 is that the contracted negative form *don’t* is restricted to the most informal written text type, social letters, while being totally absent from both student essays and press news reports. Negativity is expressed with the full form *do not* in written texts of medium and high formality. Non-agreement is most frequent in informal social letters, where it occurs exclusively in contracted forms (27.8% of all occurrences of *do* in social letters, or 5 tokens). This contracted form has above been
identified as an overt creolism which is occasionally used in informal language but, due to its salience, is avoided in written texts of higher formality. In student essays, non-agreement is rather infrequent, occurring only in 4% of the instances of *do* in this category (1 token), and it is completely absent from highly formal press news reports.

Non-agreement with *do* in the written data occurs only in sentences with negative polarity, while this is not the case in the spoken text categories. As in spoken language, non-agreement with *do* in written language decreases with increasing formality and non-agreement with the full form is restricted to the text category of medium formality.

Figure 15 (a-c) shows the most frequent verbs in the written text categories representing different relative levels of formality.

**Figure 15a: Most frequent verbs in social letters**

**Figure 15b: Most frequent verbs in student essays**
As Figure 15 a-c shows, non-agreement rates decrease with increasing formality. In social letters represented by Figure 15a, there is some non-agreement with most of the verbs, ranging between 5.7% with the verb *have* and 13.3% with *say*. The negative contracted *don’t* is the verb with most non-agreement, namely 62.5%. In student essays, non-agreement rates are generally slightly lower, between 5.6% for *(not) do*, i.e. the full negative form of *do*, and 7.1% for *follow*. In this text category, the verb *make*, with 15.4%, has the highest rate of non-agreement. There is little non-agreement at all in the text category of press news reports, and this is mostly with the verb *say* (7.7% of all occurrences of this verb are instances of non-agreement). Thus formality is a factor which strongly influences non-agreement also in written language.

The most frequent verb in all written text categories is *have*, as the total numbers are considered irrespective of whether it is used as full verb, semi-auxiliary or auxiliary. The rates of non-agreement with *have* are similar in social letters and student essays (5.7% and 6.5% respectively) and considerably lower in press news reports (1.9%).

With respect to agreement rates, the two less formal text categories resemble one another, but the formal press news reports can be clearly distinguished, as the non-agreement rate is considerably lower than in the former two text types.

The same difference between the two less formal versus the most formal category is found when the five most frequent verbs (besides *have*) are taken into consideration. In social letters, the most frequent verbs are *seem, say, want, work* and *don’t*, while *(not) do, need, follow, allow* and *make* occur most frequently in student essays and *say, (not) do, understand, represent* and *continue* are most frequent in press news reports. There is an
obvious difference in the degree of formality between the verbs in press news reports and the two less formal text categories, but less of a difference between the latter two text types.

Many of the most frequent verbs in press news reports are used in recurring expressions such as ‘The Gleaner understands…,’ which correspond to similar expressions found frequently in my spoken data, such as ‘RJR News understands that…,’ etc. The verbs in the written news category, however, are less focused on reporting and centre more strongly on an evaluation of events than in the corresponding spoken texts. In contrast to news broadcasts, the list of most frequent verbs in press news reports contains fewer verbs of reporting, such as *report, state* or *note*, but a higher number of evaluative verbs such as *represent* used in expressions such as *This represents a big increase over the $1.5 million paid out in 1991* (W2C-005). Although differences between the categories are visible, the Figures 15 a-c do not reveal as strong a topical preference in any of the written categories as Figures 7 a-c do in the spoken text types. Yet the above discussion shows that both the frequency of non-agreement and the frequency of particular lexical items is strongly tied to register characteristics and the level of formality.

Statistical tests\(^{85}\) show that, in the written data under consideration, there is indeed a significant correlation between the level of formality of a certain text category and the amount of non-agreement with lexical verbs. The corresponding Chi square test \(\chi^2 = 28.577, \text{df} = 2, p= 0.000\) is highly significant, and a logistic regression in which grammatical non-agreement is the dependent variable and formality is the independent variable yields the result presented in Table 10.

<table>
<thead>
<tr>
<th>Variables in the Equation</th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>df</th>
<th>Sig.</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1(^{a}) Formality</td>
<td>-0.635</td>
<td>0.127</td>
<td>25.196</td>
<td>1</td>
<td>0.000</td>
<td>0.530</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.301</td>
<td>0.251</td>
<td>26.777</td>
<td>1</td>
<td>0.000</td>
<td>0.272</td>
</tr>
</tbody>
</table>

\(^{a}\) Variable(s) entered on step 1: Formality.

**Table 10: Logistic regression for grammatical non-agreement and formality (written)**

The correlation between grammatical non-agreement and level of formality is negative, as indicated by the variable B (B = - 0.635 in step 1). This means that when formality levels rise there is a corresponding decrease in non-agreement. The same tendency was found in

\(^{85}\) Cf. footnote 73. The results for the verb *be* in the written data are discussed in the next section.
spoken language, but the correlation there was even stronger. Formality is thus a good predictor of non-agreement not only in spoken but also in written language, and this is confirmed by the highly significant Wald test (Sig. = 0.000).

The result of a Chi square test for subject number ($\chi^2 = 0.679$, df=1, p=0.409) shows that this is a factor that does not affect non-agreement rates with lexical verbs in any significant way across the written data, and this result is also confirmed by a non-significant Wald test (Wald = 0.678, Sig.= 0.410). This means that it is not possible to make predictions about the amount of non-agreement to be expected in written language based solely on whether the subject is singular or plural. This result is most likely due to the high number of hypercorrectly marked verbs in student essays and, to some extent, in press news reports.

In order to test whether the potential explanations for non-agreement described above, such as simple unmarking of verbs, notional considerations or the proximity of a noun other than the controller of agreement, are tied to formality or subject number in any statistically significant way, a cross-tabulation was carried out. The result for potential explanation versus formality yields a Chi square approximation of $\chi^2 = 38.536$.\(^{86}\) This indicates that a correlation indeed exists between formality and the described causes of non-agreement. The contingency coefficient ($C = 0.504$) points to a moderately strong correlation.

The same test of dependence of the potential causes of non-agreement on subject number yields an approximation of $\chi^2 = 46.068$ (df = 6), a result that is highly significant on the 0.000 level. The contingency coefficient ($C = 0.538$) indicates that the correlation between subject number and type of non-agreement is moderately strong. While subject number does not influence the amount of non-agreement, it does have an effect on potential causes of non-agreement in that, of course, simple unmarking of verbs, the single most frequent explanation for non-agreement, occurs only in sentences with singular subjects, while only plural subjects can lead to hypercorrect addition of inflectional marking.

In the following section, subject-verb agreement with be will be analysed in the three written text categories representing low, medium and high formality levels. The types of non-agreement and the dependence of non-agreement on subject number and formality will

\(^{86}\) Some of the frequencies in my data are too low for a reliable Chi square computation, as this statistical test relies on a minimum expected frequency of 5 instances in each category. SPSS therefore computes only a Chi square approximation, which, however exceeds the critical value of 28.299 at df = 12 and is still highly significant. While this result is based on only an approximation and thus cannot be assumed to be reliable, it does tentatively support the assumptions that the correlations described above are a stable agreement pattern in Jamaican English. In order to obtain a more reliable result, the contingency coefficient has to be taken into account, too. Cf. also footnotes 74 and 75.
be in the centre of attention, but tense as an additional potential influence will also be considered.

**4.4.2 Agreement with the Verb BE**

As Table 7 (in section 4.3.2) shows, there are in total 3155 instances of *be* in my written data, 744 of which were found in social letters (23.6% of all verbs in the written data), 1392 in student essays (44.1%) and 1019 in press news reports (32.3%). Social letters, the most informal category, has the lowest number of occurrences, probably due to the smaller number of texts in this category. Table 11 shows the distribution of agreement and non-agreement with *be* across the three written registers.

<table>
<thead>
<tr>
<th>Formality</th>
<th>Agree</th>
<th>Non-agr</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tokens</td>
<td>%</td>
<td>Tokens</td>
</tr>
<tr>
<td>Low</td>
<td>737</td>
<td>99.1</td>
<td>7</td>
</tr>
<tr>
<td>Medium</td>
<td>1350</td>
<td>97.0</td>
<td>42</td>
</tr>
<tr>
<td>High</td>
<td>1004</td>
<td>98.5</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>3091</td>
<td>98.0</td>
<td>64</td>
</tr>
</tbody>
</table>

Table 11: Distribution of agreement and non-agreement with *be* in the written registers

The overall amount of non-agreement with the verb *be* is very low, only 2% of present tense occurrences of *be* do not agree with their subjects. The rate is lowest in social letters, (0.9%), and highest in student essays (3%), with press news reports ranging in between (1.5%). Non-agreement rates with *be* are negligible in the written data, but an analysis may reveal interesting results if compared to the patterns found in the spoken data (for the comparison see section 4.5). A Chi square test (χ² = 12.864, df = 2, p = 0.002) reveals that formality exerts a significant influence on the rates of non-agreement with *be* in my written data. However, the correlation between formality and the rates of non-agreement is not a linear one. In contrast to lexical verbs, there is no clear tendency of overall non-agreement rates to decrease with increasing formality. Instead, non-agreement is least frequent in the most informal text type, and most frequent in the text category of medium formality, with the highly formal text type ranging in between. A thorough analysis of the data with regard to
specific types of non-agreement is therefore necessary in order to explain why this pattern emerges and in which way it is determined by effects of formality. 

As concerns the grammatical number of the subject, in all written text categories under consideration there is more non-agreement when the subject is plural than when it is singular. This is in stark contrast to lexical verbs, where most non-agreement occurred in sentences with singular subjects, but confirms the results obtained in the spoken data (cf. section 4.3.2). Non-agreement with singular subjects rises from 0.3% in social letters to 1.9% in student essays and decreases to 0.8% in press news reports, while non-agreement with plural subjects is almost equally frequent in social letters and student essays (5.0% and 5.3% respectively) and decreases down to 2.9% in press news reports. The Chi square test for subject number ($\chi^2 = 33.077$, df = 1, p = 0.000) reveals that this is a factor that significantly influences non-agreement rates with *be* in written Jamaican English.

Turning to the potential effects of tense on non-agreement we find that non-agreement rates do not differ systematically when the sentence is in present rather than past tense and vice versa. In social letters, 0.8% of the verbs in present tense do not agree with their subjects, while 1.6% of past tense verbs are cases of non-agreement. In student essays, non-agreement rates are higher in both present and past tense verbs (2.6% versus 5.1% respectively), and the rates decrease in press news reports, in which 1.6% of sentences with present tense reference show non-agreement as compared to 1.4% with past tense reference. The rates do not differ systematically by tense. Thus, tense is not a significant influencing factor on non-agreement rates in my written data, and this is confirmed by a non-significant Chi square test ($\chi^2 = 0.455$, df = 1, p=0.499).

Despite the small rates of non-agreement, the factor subject number appears to be the strongest predictor of non-agreement in my written data, followed by formality as the second strongest indicator. This contrasts with the spoken data, where subject number and tense were found to be influencing non-agreement and formality was insignificant.

Although the overall non-agreement rates with *be* are negligible, the types of non-agreement occurring in the written data will also be analysed, in order to test whether the overall pattern is similar to that found for lexical verbs or whether there are crucial differences. Figure 16 gives an overview of the types of non-agreement found in each register. The percentages refer to the proportion of the total non-agreement within the respective text category.
The category ‘Verb number’ refers to all instances of non-agreement with *be* that cannot be explained by notional considerations, the proximity of a local noun, the presence of existential constructions, or ‘other’ types of non-agreement but which constitute in a simple mismatch of number between the verb form and its subject. These non-agreements are least frequent in social letters (14.3%) and most frequent in student essays (26.2%), with press news reports ranging in between (20.0%) and thus this category neatly mirrors the overall frequency of non-agreement across the written registers. In the majority of cases the subject number is singular and the verb used is plural, but occasionally a plural subject is followed by a singular verb.

The previously discussed pattern of notional agreement is, again, shown to be stable, in that notional agreement is highest in the most formal text category and accounts for much of the non-agreement in this category (40.0%). Similar results have been found for both modes of production, spoken and written. This tendency also does not depend on the type of verb, whether lexical or the exceptional verb *be*. That notional agreement appears to be higher in social letters than student essays is caused by the very low number of non-agreement in the former category (only 7 tokens in total) as compared to the much higher number of instances in the latter text type (42 tokens). The one instance of notional agreement in social letters accounts for 14.3% of all non-agreement in this category, while the 5 instances in student essays account for only 11.9% of the total non-agreement in this text type.
None of the instances of notional agreement in the written data are caused by the plural interpretation of an indefinite subject pronoun. In most cases, a collective noun is involved or an expression similar to that in example (58).

(58) …, he disclosed that during that year **96,000 pounds of ganja was** discovered.
    (W2C-007)

Here, the noun phrase **96,000 pounds of ganja** is interpreted as a total amount rather than a number of individual entities and is thus followed by a singular verb.

Proximity agreement is the prevailing type of non-agreement in all written registers considered. Proximity agreement accounts for 28.6% of all non-agreement in the text category of social letters, 42.9% in student essays and 40.0% in press news reports. Following the overall agreement pattern in my written data of the verb **be**, proximity agreement is least frequent in social letters, most frequent in medium formality student essays and of intermediate frequency in the highly formal press news reports. Proximity effects account for most of the non-agreement in the text category of press news reports, but the percentage is low when the occurrences are measured against the number of sentences containing complex noun phrases, much lower, in fact, than in the educational texts. This suggests that proximity agreement is a form which is stigmatized and avoided in published texts but due to its low salience is not even noticed by students in their university assignments.

Non-agreement in existential **there + be** constructions is so rare in my written data as to be negligible. In a total of 136 occurrences of this construction in my whole written data only three sentences were found in which the verb form does not agree with the post-verbal subject noun phrase. One of these occurred in social letters (59), where a total of 27 existential constructions were counted (this one case thus represents 3.7% non-agreement), and two cases appear in student essays (out of a total of 29 tokens of **there + be**, i.e. 6.9%), (60) and (61).

(59) The problem that we face here is that **there is limited reading materials** on your island... (W1B-008)
(60) Consequently **there is not enough funds** to finance government programmes... (W1A-002)
(61) The manual system also requires one to go to different persons to approve individual courses, since there were no system of checking pre-requisite, therefore (W1A-015)

All of these instances occur in full rather than contracted forms, but since in written language generally more full forms and fewer contractions tend to be used, this result is not surprising. None of the 80 existential constructions in press news reports contains any non-agreement at all. While the absolute numbers of such constructions increase with increasing formality, the amount of non-agreement does not. However, due to the low numbers, a meaningful analysis of non-agreement in existential there + be constructions across the written registers is impossible. A more detailed analysis of this construction will follow in chapter 5, which compares Jamaican English with four other native and non-native varieties of English with regard to existential there + be constructions. This analysis is, however, based exclusively on spoken data.

Last but not least, the category ‘other’ consists of instances that cannot be subsumed under any other category in the figure. While in the written texts of low formality, both cases of ‘other’ non-agreement (adding up to 28.6% of all non-agreement in this category) are the result of coordinated subjects leading to non-standard agreement, one of which is presented here as (62), three of the six cases of ‘other’ non-agreement in student essays (which account for 14.3% of the non-agreement) can be so explained.

(62) By the way Mr. and Miss Preston was extremely good. (W1B-006)

The remaining three instances of non-agreement are caused by noun phrase-internal non-agreement, that is, they are the result of a lack of plural marking in the noun, as shown in (63).

(63) These feature are peculiar to passive movement. (W1A-017)

In press news reports, no ‘other’ non-agreement was found. The patterns found confirm that the verb be is exceptional in many respects and functions differently from the lexical verbs discussed above.
4.4.3 Conclusion – Subject-Verb Agreement in Written Jamaican English

Although the total amount of non-agreement in written Jamaican English is rather low, the analysis of subject-verb agreement across the three registers social letters, student essays, and press news reports, representing low, medium and high levels of formality respectively, has revealed some internal variation. This variation is not random, but a number of linguistic and non-linguistic factors have been found to systematically influence grammatical agreement and non-agreement in the data. The tendencies found are as follows:

First, subject number does not influence the rate of non-agreement with lexical verbs to any significant extent, but it is the best predictor of non-agreement with the exceptional verb be. In all registers, there is more non-agreement with be when the subject is plural than when it is singular. As described in section 4.3.2, this finding can be attributed partly to the invariant copula iz/woz in the mesolect, which indirectly influences the use of copular verbs in Jamaican English by facilitating an extension of is and was to plural contexts, and to notional considerations.

With lexical verbs, there is a difference in non-agreement rates only in the most informal text category of social letters, where non-agreement is much higher with singular than plural subjects. This difference in non-agreement rates in sentences with singular as opposed to plural subjects is not evident in student essays and press news reports, where the rates are almost equal for both numbers. This pattern is the result of the high number of unmarked verbs following singular subjects in social letters which coincides with an almost complete absence of hypercorrectly marked plural verbs in this category. In student essays and even more in press news reports, the rates of unmarked verbs decrease notably, while hypercorrections increase substantially, thus the difference in non-agreement rates between sentences with singular and plural subjects almost disappears. It is, however, not subject number that actually causes this pattern, but an interaction has been found between subject number and formality.

Second, formality, the only non-linguistic variable considered here, is a very good predictor of non-agreement rates with lexical verbs, but a less useful one with the verb be. With lexical verbs there is very straightforwardly a negative correlation between formality and overall rates of non-agreement, in that the amount of non-agreement is highest in the
most informal text category social letters (10.6%) and decreases with increasing formality to 8.5% in student essays and down to 2.7% in press news reports. We can therefore postulate a stylistic continuum ranging from most non-standard in informal texts, both with regard to the rate of non-agreement and the range of variants used, to most standard in formal language. What has to be kept in mind here, however, is that the two less formal text types are not as homogeneous as the most formal category, because the category social letters includes very informal private emails as well as more formal business letters, while the student essays comprise essays written in examination situations as well as assignments produced under less formal circumstances (cf. section 4.2.2 on the registers chosen for this study). This is reflected in the analysis, too, in that the two categories tend to resemble one another in the agreement patterns found, and the decrease of non-agreement between student essays and press news reports is much steeper than between social letters and student essays. Formality, which is closely tied to the respective characteristics of the specific register, has been found to also play a statistically significant role in determining agreement patterns with *be* in the written registers. It is, however, not entirely clear why the text category of medium formality, student essays, displays the highest frequency of non-agreement with *be*, and the most informal category, social letters, has the lowest rate, with high formality texts ranging in between.

Third, the type of verb has been proven to be important, as *be* functions differently from lexical verbs in many respects. Not only is there considerably less non-agreement with *be* than with lexical verbs, but the linguistic and non-linguistic constraints also do not affect lexical verbs and *be* to the same extent, nor in the same way, as illustrated above. Tense has not been analysed for lexical verbs, and it turned out not to be a factor which influences non-agreement with *be* to any significant extent.

Fourth, the distribution of the different types of non-agreement proved to be rather stable. In other words, the registers correlate differently with different types of non-agreement, but these correlations found in written language are not very different from those reported earlier for spoken language. While overall rates of non-agreement decrease with increasing formality, the trend is not uni-directional, as not all types of non-agreement have decreasing frequencies in texts types of higher relative levels of formality. The fact that the lack of inflectional marking of verbs, the prevailing type of non-agreement in all registers, decreases with increasing formality, together with a strong prevalence of hypercorrect
markings in educational written texts proves that inflectional marking in verbs is a prestigious form whereas the lack of such marking is stigmatized.

This clear negative correlation between formality and non-agreement is not revealed by the data with regard to notional agreement, which increases with increasing formality. In the written data, the postulated tendency in informal Jamaican English to extend notional agreement (cf. Sand 1999) was not found, as there are no instances of plural interpretation of indefinite pronouns and extensions of notional agreement to quantifiers such as much to plural count nouns are very rare. Other instances of direct creole influence are also very infrequent, as, for example, uses of him as subject pronoun or the extension of negative contracted don’t to sentences with singular subjects. In fact, this occasional extension of don’t was restricted exclusively to the most informal text category of social letter. Thus no evidence at all was found in the written data for a general tendency toward simplification in the use of the verb do.

As concerns proximity agreement, there is no clear uni-directional tendency that is related to formality or register characteristics, as this type of non-agreement is completely absent from informal written texts, rather infrequent in highly formal news texts, and very frequent in texts of medium formality from the educational context of student essay writing.

On the basis of the discussion above the continuum postulated by Deuber “from more English to more Creole ways of speaking” (2009a: 47ff.) can be extended. First, it applies to the whole spectrum of spoken Jamaican English, and, to a lesser extent, to written Jamaican English. That is, the overall rate of non-agreement and the range of stylistic variants is greater in informal than formal usage and decreases with increasing formality, and this is true for spoken and written language alike. This is because in informal language, standard, non-standard and creole variants are available to speakers as stylistic alternatives, whereas in more formal text types the choice of variants is more restricted.

Second, Deuber introduces the notion of a hierarchy among creole forms “which probably corresponds to a large extent to their positions on the sociolinguistic continuum of varieties but seems to be determined as well by their formal salience in relation to English and the stereotypes speakers hold of them” (2009a: 29ff.). The thorough analysis of subject-verb agreement patterns in Jamaican English has revealed that it would be reasonable to extend this hierarchy and to postulate a ‘continuum of salience’ from highly salient and
stigmatized overt creolisms at one end over less salient zero forms to indirect contact influences, which are of so little salience as to go largely unnoticed even in texts of high formality. Examples of overt creolisms are the use of creole pronouns, e.g. him or them, as subject or possessive pronouns in English Deuber (2009a: 28-30), but also the extension of invariant don’t or the invariant copula is and was to plural contexts, e.g. they is, we was. These are most frequent in informal spoken domains and decrease sharply with increasing formality, and in written language they are restricted exclusively to the most informal registers. Zero forms, such as the absence of inflection in third person singular verbs, occur in all text types, spoken and written, but tend to be avoided in formal registers. Last, indirect contact influences, such as hypercorrect inflection in third person plural verbs in accordance with the proximity principle, occur mainly in texts of higher formality. While they are edited out of printed texts these forms are of so little salience as to go unnoticed in the educational sector, where they were attested most frequently.

The next section will briefly explore the mode of production as a factor that potentially influences agreement and non-agreement in educated Jamaican English.

### 4.5 Spoken and Written English Compared

The central question to be answered in this section is the following. How good a predictor of grammatical non-agreement in Jamaican English is the mode of production, i.e. whether a text was spoken or written. Not just the overall rates of non-agreement are important here but especially the types of non-agreement and possibly existing differences in the overall agreement patterns. Table 12 shows the distribution of agreement and non-agreement with lexical verbs across the three levels of formality in the spoken and written data set.

| Formality | Spoken | | | Written | | |
|-----------|--------|--------|--------|--------|--------|
|           | Tokens | %      | Tokens | %      | Tokens | %      | Tokens | %      | Total |
| Low       | 588    | 82.5   | 125    | 17.5   | 336    | 89.4   | 40     | 10.6   | 1089  |
| Medium    | 586    | 91.8   | 52     | 8.2    | 942    | 91.5   | 88     | 8.5    | 1668  |
| High      | 679    | 96.4   | 25     | 3.6    | 610    | 97.3   | 17     | 2.7    | 1331  |
| Total     | 1853   | 90.2   | 202    | 9.8    | 1888   | 92.9   | 145    | 7.1    | 4088  |

Table 12: Agreement and non-agreement with lexical verbs in spoken and written Jamaican English

---

87 This is due to the higher lexical density and the corresponding higher number of complex noun phrases in highly formal text categories.
In a direct comparison of non-agreement rates in spoken and written text categories of low, medium and high formality, the most striking difference in token numbers is found between informal conversations (125 cases of non-agreement) versus social letters (40 cases). The difference is not so great for class lessons (52 tokens) versus student essays (88 tokens) and broadcast news (25 instances) versus press news reports (17 instances). Note that numbers decline steadily with increasing formality in spoken language but not in the written data, where most non-agreement is found in the educational context of student essays.

If these numbers are measured against the overall counts of lexical verbs, the total rate of non-agreement in the written data is slightly lower (7.2%) than that in the spoken data (9.8%). In both the spoken and the written data, there is a very straightforward agreement pattern in that agreement rates decrease with increasing formality. In texts of low formality, the rate of grammatical non-agreement is higher in the spoken data, with 17.5% non-agreement, than in the written data, where there is only 10.6% non-agreement. The same is true for texts of high formality, as 3.6% of the verbs in broadcast news do not agree with their subjects in comparison to 2.7% in press news reports. In texts of medium formality, i.e. the spoken category of class lessons and the written student essays, the rate of non-agreement is similar in spoken and written language, namely 8.2% versus 8.5% respectively. The mode of production can thus be said to influence rates of non-agreement in educated Jamaican English. A Chi square test for the mode of production versus non-agreement rates across the whole data set ($\chi^2 = 9.572$, df = 1, $p = 0.002$) shows that the mode of production is indeed a significant influencing factor on subject-verb agreement in Jamaican English. However, if a Chi square test is computed for each level of formality individually, the result changes considerably. For low formality, the test yields $\chi^2 = 9.099$, a result that is significant at the 0.005 level, but in texts of medium formality, $\chi^2 = 0.079$, i.e. there is no significant influence of the mode of production on non-agreement rates. Similarly, in high formality text categories the result, $\chi^2 = 0.765$, is also not significant. This means that it is only in the most informal text category that a significant influence of the mode of production can be detected.

This is not a surprising result if we assume that agreement patterns in Jamaican English are affected by substrate influence. The work described above and previous research (cf. chapter 3) indicates that Jamaican Creole influences prevail particularly in informal spoken

---

88 In each case, df = 1.
language, causing in such informal spoken registers high numbers of non-agreement due to the absence of inflectional markings in verbs following singular subjects, as well as some other types of non-agreement (cf. section 4.3.1. above). Such creole influence is much less obvious in informal written language and largely absent from more formal spoken and written text types.

Table 13 shows the distribution of agreement and non-agreement with the exceptional verb be across the three formality levels in the spoken and written data.

| Formality | Spoken | | | Written | | |
|-----------|--------|--------|---|--------|--------|
|           | Tokens | %  | Tokens | %  | Tokens | %  | Tokens | %  | Total |
| Low       | 1507   | 99.2 | 12    | 0.8 | 737    | 99.1 | 7    | 0.9 | 2263  |
| Medium    | 1652   | 99.1 | 15    | 0.9 | 1350   | 97.0 | 42   | 3.0 | 3059  |
| High      | 1129   | 98.4 | 18    | 1.6 | 1004   | 98.5 | 15   | 1.5 | 2166  |
| Total     | 4288   | 99.0 | 45    | 1.0 | 3091   | 98.0 | 64   | 2.0 | 7488  |

Table 13: Agreement and non-agreement with be in spoken and written Jamaican English

In total, there is slightly more non-agreement in the written data (2.0%) than in the spoken data (1.0%). A Chi square test ($\chi^2 = 12.473$, $p = 0.000$) reveals that the correlation between the mode of production and non-agreement rates with be is significant. If the test is computed for each level of formality separately, the results prove that this correlation does not exist across all formality levels. For the most informal text categories $\chi^2 = 0.137$, which does not exceed the critical value of 2.706 at the $p = 0.1$ level and is thus not significant. In other words, the mode of production is irrelevant for determining non-agreement rates with be in informal Jamaican English, and this is a stark contrast to agreement with lexical verbs. This contrast is caused by the fact that –s inflection, which accounts for most of the non-agreement with lexical verbs, does not apply at all to the verb be. In the most formal text categories, the Chi square test for mode of production yields $\chi^2 = 0.034$ and is thus not significant. The text types of medium formality are the only categories where the correlation between non-agreement with be and the mode of production is highly significant ($\chi^2 = 18.599$, $p = 0.000$). This higher rate of non-agreement in written than spoken language of medium formality is above all a result of the much more frequent occurrences of proximity agreement in student essays (18 tokens) than class lessons (1 token).
In sum, the mode of production is not as good a predictor of non-agreement rates in educated Jamaican English as formality, a factor which is much better capable of explaining agreement variation in the data than is the mode of production.

With regard to subject number we found that the Chi square test is not significant for all of the six text categories, e.g. it does not exert any notable influence on non-agreement rates in class lessons, student essays and press news reports. In the other three categories, i.e. direct conversations, social letters and broadcast news, subject number seems to influence the agreement pattern. However, there is a strong interaction with formality, a factor that causes the characteristic pattern of non-agreement, as absence of –s marking in verbs following singular subjects is restricted to informal text types while hypercorrect markings of verbs following plural subjects occurs mainly in texts of medium and high formality. Therefore, subject number cannot be considered a good predictor of non-agreement in educated Jamaican English.

Let us now turn to a comparison of the prevailing types of non-agreement with lexical verbs occurring in the spoken and written text categories of low, medium and high formality as shown in Figure 17.

![Figure 17: Types of non-agreement with lexical verbs in spoken and written Jamaican English](image)

Obviously, absent –s marking is the most prevailing type of non-agreement, accounting for more than half of all non-agreement in all text categories, irrespective of the formality level or the mode of production. That means that non-agreement with lexical verbs mostly consists in a lack of inflectional endings in locations where they would be expected to

---

89 This comparative analysis of types of non-agreement is based on the results for lexical verbs only.
occur in the standard variety rather than the hypercorrect addition of inflection, and this is true for spoken and written language use alike. Although the presence of these zero forms in all text categories suggests that it might be a general simplification tendency in Jamaican English, their prevalence in informal texts points to contact influence as a potential explanatory factor, as lack of inflectional marking is a characteristic feature of Jamaican Creole which facilitates its occurrence in informal Jamaican English. However, zero marking is not very salient in Jamaican English, as noted in section 4.4.3. Differences between the modes of production can be found particularly with regard to hypercorrections and non-agreement with the contracted negative don’t. Hypercorrections are much more prevalent in the written texts than the spoken ones, and more so in formal than informal text types. The higher frequency of hypercorrect inflectional endings on verbs following plural subjects in texts of medium and high formality than in texts of low formality, together with the fact that such hypercorrections are most prevalent in the educational context of student essays, suggests that –s inflection is a prestigious form, while the absence of such inflection is stigmatized and thus tends to decrease with increasing formality.

Although don’t non-agreement occurs in both spoken and written language, a distributional difference has been found, in that in written language it is restricted to the most informal text type, while in spoken language it is found in texts of low and medium formality.

In order to compare agreement patterns in spoken and written language more thoroughly, one type of non-agreement with be discussed in detail. As the single most frequent type of non-agreement present in my data is the simple absence of –s marking in lexical verbs following singular subjects, the analysis will be based on the frequencies and potential explanatory factors of this type of non-agreement. Absent –s is indeed very frequent in both modes of production, but in spoken language it is even more prevailing among the types of non-agreement found in the data (163 tokens, or 80.7% of all non-agreement) than in written language (86 tokens, or 59.3% non-agreement). My result thus confirms those by Deuber (2009a) who found that, in her informal spoken data, “simple unmarking of third person singular verbs is clearly the predominant phenomenon.”

In the spoken text categories absent -s occurs 102 times in direct conversations (81.6% of all non-agreement in this text category), 43 times in class lessons (82.7%) and 18 times in broadcast news (72.0%). In the written registers it is found 29 times in social letters (72.5%), 48 times in student essays (54.6%) and 9 times in press news reports (52.9%).
other words, in informal and highly formal texts, the token numbers of non-agreement are higher in the spoken than in the written data. Only the text categories of medium formality deviate from this pattern, and here the difference in absolute numbers is rather small, (43 versus 48 tokens), but the proportions of the total amount of non-agreement accounted for by this type of non-agreement differ considerably, as absent –s adds up to 82.7% in class lessons and only 54.6% in student essays. What is more relevant than absolute token numbers, however, is the pattern of factors that cause this absence of inflection in singular verbs, and it is this pattern that will be the focus in the following analysis.

Figure 18 shows the proportion of grammatical non-agreement caused by the absence of –s inflection that can be attributed to particular explanatory factors in the spoken and written registers of low, medium and high formality.

In terms of total numbers, the number of non-agreement involving simple unmarking is always higher in the spoken register than in the corresponding written register of similar formality, i.e. considerably higher in direct conversations (69 tokens) than social letters (23 tokens), higher in class lessons (38 tokens) than student essays (21 tokens) and higher in broadcast news (5 tokens) than press news reports (1 token). The percentages of the total amount of non-agreement involving absent inflection, however, do not display this straightforward trend. While there seems to be a stable trend in written language of rates of unmarked verbs to decrease with increasing formality, from 79.3% in social letters down to 43.8% in student essays and down to 11.1% in press news reports, this trend does not appear in the spoken data, where the rate rises from 67.6% in direct conversations to 88.4% in class lessons but then decreases sharply to 27.8% in broadcast news. This can be best explained when other potential explanations for the absence of –s marking are considered.
An obvious difference between spoken and written language exists in the amount of non-agreement co-occurring in environments where other overt creole features are found. Such environments appear exclusively in the most informal text categories, while being totally absent from texts of medium and high formality. This indicates that these direct creolisms are highly salient and stigmatized, and therefore tend to be avoided in formal situations (cf. section 4.4.3). This type of non-agreement is rather frequent in the most informal spoken category (15 tokens, or 14.7%) but in the corresponding written category there is only one instance (3.5%). Thus direct contact with Jamaican Creole is largely restricted to spoken language, and my results thus confirm Mair’s postulation that “in view of the ever present role of the creole in speech, it is surprising to note how limited a role it plays in writing” (2002: 55).

Non-agreement that is due to a noun phrase-internal mismatch of number, i.e. involving the lack of plural marking in nouns, is more frequent in written than spoken language and occurs most often in student essays, possibly due to the high complexity of noun phrases in this category which may be assumed to result in occasional agreement errors by L2 speakers. However, as this type of non-agreement cannot always be attributed to non-standard (lack of) verbal inflection but is sometimes caused by the absence of plural marking on nouns following numerals or other quantifiers, it will not be analysed in detail here and further research will have to be carried out to shed more light on this phenomenon.

The results show that proximity agreement occurs more frequently in my written data (27 tokens in the case of lexical verbs) than in the spoken data (7 tokens). My spoken data seem to support Francis’ view that “proximity concord is occurring in the upper stylistic registers” (1986: 316), since the token numbers rise steadily with increasing formality. However, the same is not true for the written data, where proximity agreement is absent from the most informal text category and most frequent in texts of medium formality, with highly formal texts ranging in between. This means that Francis’ postulation that examples of proximity agreement derive “mostly from edited English” (1986: 316) cannot be supported on the basis of my data. Francis’ work, however, was based on a casual collection of examples in a number of registers rather than on systematic corpus investigation. What Francis did not consider, but what is in fact more important than the

---

90 For a thorough analysis of the frequency and functions of creole features in written (computer-mediated) Jamaican English see Hinrichs (2006).
percentages of non-agreement within particular text categories that can be explained by proximity effects, are the percentages of non-agreement within all sentences containing complex noun phrases, i.e. those where proximity could possibly occur. Table 14 gives the numbers and respective percentages of complex noun phrases and instances of non-agreement across the spoken and written text categories.

<table>
<thead>
<tr>
<th>Formality</th>
<th>Spoken</th>
<th>Written</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>Complex NP</td>
<td>27</td>
<td>53</td>
</tr>
<tr>
<td>Proximity Agr.</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Proximity %</td>
<td>7.41%</td>
<td>3.77%</td>
</tr>
</tbody>
</table>

Table 14: Complex noun phrases and proximity agreement in spoken and written Jamaican English

Table 14 indicates that Francis’ postulation of proximity agreement generally occurring more frequently in texts of “upper stylistic registers” (Francis 1986: 316) is only partly true for my Jamaican data, being least frequent in informal text categories but most frequent in texts of medium rather than high formality. The postulation that proximity effects can be observed mainly in edited texts is not confirmed, as the phenomenon is rather infrequent in the category of press news reports.

While it is true that there is more proximity agreement in written than spoken language in terms of absolute frequencies, this most likely results from the general tendency for written language to contain more complex or ‘heavy’ subject noun phrases than spoken texts of similar formality. Except for my informal registers, where the number of complex noun phrases is equal (27 in conversations and 26 in social letters), this tendency is very obvious in the registers of medium formality, with 187 tokens in the written versus 53 tokens in the spoken register, and in the registers of high formality, where 138 complex noun phrases were found in press news reports versus 97 in broadcast news.

Quirk et al.’s (1985) postulation that “[p]roximity concord occurs mainly in unplanned discourse” is confirmed by my results for the spoken data set, but it does not hold in the case of my written data, as proximity effects are completely absent from the most ‘unplanned’ and spontaneous of my written text types, social letters. The second of Quirk et al.’s statements, that “[i]n writing it will be corrected to grammatical concord if it is noticed” is confirmed for the carefully edited published register of press news reports, but
not for the category of student essays, the register with the highest amount of proximity agreement.

Although Francis’ and Quirk et al.’s views seem to contradict one another, evidence for both can be found in the data, depending however on whether the analysis is based on absolute token numbers or on the percentage of complex noun phrase subjects leading to proximity effects as compared to the total number of such noun phrases where proximity agreement would be possible.

Last but not least, there are crucial differences in spoken versus written language as concerns notional agreement. In both modes of production, notional agreement is most prevalent in texts of highest formality, the category where most collective nouns occur. The fact that collective nouns are most frequent in highly formal texts again provides evidence for the influence of register characteristics on the subject-verb agreement pattern.

In the written data there is a straightforward trend for notional agreement to increase with increasing formality, but the same cannot be observed for the spoken language, where there is a high amount of notional agreement in informal language, then the rate decreases sharply in texts of medium formality only to rise again in formal texts. The discussion in section 4.3.1 indicates that the general trend of an increasing influence of notional considerations with increasing formality is overruled in the informal spoken text category by an extension of notional agreement to indefinite pronouns such as everybody, but this phenomenon is not observed in the informal written texts. This apparent effect of the mode of production on notional agreement in informal language use is to be attributed to the fact that the extension of notional agreement to indefinite pronouns is a creole feature, and the discussion above showed that substrate influence is largely restricted to spoken language.

My results support Hundt’s (2006) postulation that substrate influence does not play a significant role in agreement with collective nouns in acrolectal English of the outer-circle varieties (cf. Kachru 1997) at least not in Jamaican English, where contact effects are found in subject-verb agreement with indefinite pronouns but not with collective nouns. As described above (section 3.1), Hundt attributes the “greater internal stylistic homogeneity” (2006: 223) in the outer-circle varieties, i.e. the strong orientation towards the written standard, to the “important role of the educational sector with (1) its emphasis of written production, (2) overt notions of “correctness,” and (3) a largely imported (written) norm as its yardstick for comparison” (Hundt 2006: 223). This strong orientation towards the
written standard and the important role of the educational sector is reflected in my results in that the prestigious or ‘standard’ form, the presence of -s inflection, is most frequently hypercorrectly added to verbs in the educational context of student essays.

Turning to subject-verb agreement with the exceptional verb be, we note that, despite the low rates of non-agreement, the factor subject number appears to be the strongest predictor of non-agreement in the spoken and written data, partly due to notional considerations and partly due to influence from the invariant mesolectal copula iz and woz facilitating the extension of is and was to plural contexts in Jamaican English. There are, however, differences relating to the mode of production with regard to the factors formality and tense. While formality is the second strongest indicator of non-agreement in the written data, it has been found to be insignificant for non-agreement in the spoken data set. The pattern is reversed for tense, which is the second strongest indicator in the spoken data set but insignificant in the written text types. The types of non-agreement that occur with be are not considerably different from those with lexical verbs.

4.6 Conclusion

In order to remedy the unfortunate lack in previous research of systematic register investigations taking into account not only the influence of the mode of production on subject-verb agreement patterns but also situational effects such as formality or the characteristics of particular registers, and the complex pattern of substrate influence on Jamaican English, this work has provided a detailed analysis of the subject-verb agreement in educated Jamaican English on the basis of the Jamaican section of the International Corpus of English. For the purpose of this study, three spoken registers differing in their relative level of formality were thoroughly analysed and the results were compared to those from three corresponding written registers. The investigation revealed that there is a substantial amount of linguistic variation, and the following conclusions can be drawn.

First, substrate influence is much easier to detect in lexical verbs (including do and have) than the verb be. The high frequency of absent third person –s in lexical verbs can be attributed to contact with Jamaican Creole, which generally lacks inflection and thus facilitates the occurrence of unmarked verbs in Jamaican English, particularly in informal registers.
By contrast, creole influence on the use of *be* is harder to explain. First, because *be* has no inflection whose absence can be attributed to contact effects, and second, because there is no uni-directional tendency with regard to rates of non-agreement across the formality levels which would support the assumption of creole associations. In the Jamaican mesolect, the basilectal copula *a* (equative), *de* (locative) or zero (adjectival predication) may alternate with invariant (but tense-marked) *iz* or *waz* (cf. Patrick 2007: 139). But the same is not true for acrolectal English, where Deuber (2009a: 14f.) finds the first two of these to be very rare and zero copula to be somewhat more frequent.91 However, in educated Jamaican English, standard forms of *be*, i.e. those agreeing with their subjects, alternate with non-standard forms, i.e those not agreeing with their subjects. The finding that non-standard agreement is more frequent with plural than singular subjects in my data can be attributed partly to an extension of *is* and *was* to plural contexts,92 and partly to other factors, such as notional considerations.

Second, overall non-agreement rates with lexical verbs tend to decrease with increasing formality, and this continuum was found both in spoken and written language. Moreover, there is a cline from a wide spectrum of variants93 in informal text types, comprising standard, non-standard and creole forms (such as verbs lacking third person singular –s or invariant *don’t* with all persons and numbers), to a much narrower range of alternatives in texts of higher formality, and a corresponding much higher proportion of standard forms, e.g. inflected verbs and agreeing negative forms of *do*.

Third, there is more non-agreement in spoken than in written texts, and in particular more types of non-agreement that can be attributed to direct substrate influence. In other words, standard English is clearly the norm in writing (cf. Hundt 2006) irrespective of the register, and creole features are only occasionally used for rhetorical effects in the more informal text types (cf. Hinrichs 2006), but in speech the data provide plenty of evidence of a continuum “from more creole to more English ways of language use” (Deuber 2009a: 47).

Fourth, the level of formality and characteristics of the register are much better capable of explaining the types of variation in the subject-verb agreement system than is the mode of production, i.e. whether a text is spoken or written. However, the mode of production does have effects on the agreement patterns, as some types of non-agreement prevail in the more

---

91 Sentences with zero copula were excluded from the present investigation. See section 4.2.3.
92 Though the subjects in these cases are rarely personal pronouns (*we, they*) and much more often full noun phrases, e.g. *her parents was upset*.
93 The fact that none of the features are categorically chosen or avoided even in informal registers, e.g. standard forms are also produced in informal texts, shows that all forms are available as stylistic variants.
carefully edited written registers, such as the tendency to hypercorrectly mark verbs following plural subject in texts of higher formality (an indirect contact influence), while others are restricted to spoken language, e.g. the types of non-agreement associated with direct substrate influence. The various effects of language contact led me to extend Deuber’s (2009a) hierarchy among creole features and postulate a ‘continuum of salience’ ranging from highly salient overt creolisms, such as invariant don’t for all persons, over less salient zero forms to largely unnoticed indirect contact effects, such as the hypercorrect proximity agreement in written educational texts.

The investigation of three spoken and three written registers has shown that the subject-verb agreement system in Jamaican English is rather complex, and that various linguistic and non-linguistic factors influence and contribute to the agreement pattern. Moreover, the systematic analysis shows that the postulated differences between spoken and written language are mainly restricted to informal text types, while the patterns in texts of higher formality are rather similar. The relative level of formality together with the characteristics of the register (especially the presence or absence of creole influence) are the factors that are best suitable for explaining the subject-verb agreement system in educated Jamaican English.
5 Agreement in Existential There + Be Constructions in a Cross-Variety Perspective

5.1 Introduction

The analysis of register variation presented in chapter 4 revealed a substantial amount of variation in the subject-verb agreement system particularly in informal spoken Jamaican English. As the discussion indicates, much of this variation can be explained by influence from Jamaican Creole, as many of the types of non-agreement found in the data seem to be caused by parallel structures in the creole. The absence of –s inflection on verbs following third person singular subjects is one example of such a creole feature, since in mesolectal creole, verbs tend to lack both past tense inflection and number inflection. The extension of the contracted negative don’t can be attributed to the mesolectal form duont (cf. Patrick 2004: 417f.), as can the extension of notional agreement to indefinite pronouns such as everybody and to the quantifier much, which, in Jamaican Creole, can be used to determine plural count nouns as well as mass nouns (cf. Sand 1999). All these features have above been associated with the Jamaican substrate language.

Effects of language contact as illustrated above might also be an influential factor in other post-colonial varieties of English, such as Indian English or Singapore English. However, as Mesthrie notes,

the large number of similarities across L2 Englishe s needs to be explained more carefully than in the past, where the default assumption has often been interference from the substrates. […] It is prima facie implausible, […] that over a thousand languages should induce the very same (or very similar) influences. (2004: 1141)

Similarly, Platt, Weber and Ho (1984: 47) argue that the absence of number inflection is a typical feature of New Englishes irrespective of the background languages. It is thus not entirely clear whether it is really influence from the background languages that is responsible for the characteristic patterns relating to these features or whether the pattern is caused by underlying universal tendencies that are present in post-colonial varieties of English generally (cf. section 2.1.3).

In order to test whether substrate influence indeed qualifies as a plausible explanation for such similarities across post-colonial Englishes, one syntactic construction will be thoroughly investigated in the three L2 varieties Jamaican English, Indian English and Singapore English, varieties which were all historically British-oriented but have substrate
languages of very different types and thus cannot be expected to be influenced in similar ways, and the results will be compared to those from the ICE-corpora of the two native varieties British and American English.\textsuperscript{94} If substrate influence indeed occurs in the three post-colonial varieties, the patterns should differ markedly not only from both native varieties, but also from one another.

The focus of attention will be on patterns of non-agreement in existential \textit{there + be} constructions, as they feature prominently in the literature on variable agreement in (mainly native) varieties of English, such as British English (e.g. Britain 2002, Martinez and Martinez 2003), Australian English (cf. among others Korhonen 2007 and Eisikovits 1991), American English (Crawford 2005, Hazen 2000 and Schilling-Estes and Wolfram 1994), Tristan da Cunha (Schreier 2002), New Zealand English (Britain and Sudbury 2002) and Canadian English\textsuperscript{95} (Hay and Schreier 2004 and Meechan and Foley 1994), they are easy to extract from corpora and have been identified as constructions with a relatively high frequency. Moreover, the analysis of this structure on the basis of the whole spoken corpus sections, i.e. samples of around 600,000 words per corpus, will hopefully provide a clearer picture of the influence of the linguistic factors tense, contractedness and polarity on agreement that has been only tentatively postulated in section 4.3.2 on the basis of a much smaller sample of roughly 120,000 words.

The chapter is organised as follows. Section two presents a brief overview of past research into existential \textit{there + be} constructions in varieties of English. This is followed by a description of the data and methods of analysis applied in this study. After that the results obtained in the analysis of agreement and non-agreement in \textit{there + be} constructions in educated Jamaican English will be presented with regard to the grammatical factors subject number, tense, contractedness and polarity, and these will be compared to the results from the relevant native and non-native varieties. The chapter will end with a discussion of the results and a short conclusion.

### 5.2 Previous Research on Existential There + Be Constructions

Variation in subject-verb agreement has been characterized as a widespread phenomenon in varieties of English around the world and features prominently in discussions of

\textsuperscript{94} As there is presently no ICE-USA, the comparison will be based on the Santa Barbara Corpus of Spoken American English. See section 5.3 for details of the data used for this study.

\textsuperscript{95} See also Walker’s (2007: 151f.) overview of previous research into variable agreement in spoken language data.
vernacular universals and angloversals (cf. sections 2.1.3). As Hay and Schreier (2004) point out, variationists have been particularly interested in subject-verb agreement with different types of plural subjects such as personal pronouns (they were), noun phrases (the cows were) and existential subjects (there were cows) (examples from Hay and Schreier 2004: 211), with many scholars concluding that “[n]onstandard –s exhibits much higher rates in existentials than in subject-verb structures” (Rupp 2005: 259). These constructions thus seem to be an excellent basis for a comparison of non-standard agreement and its potential causes in native and non-native varieties of English.

Existential constructions are commonly used “to focus on the existence or occurrence of something (including the non-existence or occurrence of something)” (Biber et al. 1999: 951). In English, this is usually done by means of sentences in which the existence of some indefinite entity is expressed, as in (1).

(1) Only in India there is manual work (ICE-IND, S1A-073)

In existential there + be constructions (ETB), there does not function as a locative adverbial or deictic element, as is the case for instance in (2), which is not of interest in the present study, but it takes the position of the syntactic subject.

(2) But it’s there for so many more people (ICE-GB, S1A-002)

It is, however, not the syntactic subject but the notional subject noun phrase in post-verbal position that controls subject-verb agreement. In other words, since the notional subject manual work is singular, the verb be also takes singular form.

The syntactic functioning of agreement in existential constructions is rather controversial and has been widely discussed in the literature, as summarized by Martinez and Martinez (2003: 263). According to the authors, some scholars, for instance Lakoff (1987), describe the relationship between the post-verbal noun phrase and the verb as a case of “indirect agreement, with there agreeing with the NP and the verb agreeing with there” (Martinez and Martinez 2003: 263, emphasis in original), while other linguists such as Quirk et al. (1985), Biber et al. (1999) and Crawford (2005) interpret it as a relationship of only two elements, namely the verb and the post-verbal noun phrase, in which the verb agrees in
number with the notional subject, “while there functions as a nominal reference item that refers to singular as well as plural nominals” (Martinez and Martinez 2003: 263). It is the second view that is adopted in this paper, i.e. not the ‘dummy’ there in syntactic subject position but the post-verbal noun phrase determines agreement. Corbett (2006: 64) considers this a case of “brother-in-law-agreement,” in that, as Schiffrin (1994: 239f.) notes, existential there + be constructions fulfil the “principle of end weight” (Crawford 2005: 42) by moving heavy subject noun phrases to the end of the clause, thus facilitating their cognitive processing. Walker (2007: 150f.) summarizes previous research and argues that different mechanisms have been found to be responsible for singular and plural agreement in there + be constructions. In this view, plural agreement is the result of specifier-head agreement, “either through identity of there with the postverbal subject or some sort of covert movement” (Walker 2007: 150) and singular agreement is the unmarked or default form which occurs whenever the subject is singular or unavailable (Walker 2007: 150).

What I call non-agreement throughout the present paper is cases in which the verb number does not correspond to the grammatical number of the post-verbal noun phrase, as illustrated with examples from ICE-Jamaica in (3) – (5).

(3) Well there is the shops (S1A-030)

(4) Because if this is your attitude other parents there's about twenty parents and the others apologise (S1A-098)

(5) I think there was four books (S1A-064)

Although non-agreement in existential there + be constructions has been found to be much more frequent with plural subjects, non-standard agreement does occur also with singular subject, albeit quite rarely. In the examples (6) and (7) from ICE-India the post-verbal noun phrase is singular but the verb is plural.

(6) There are some waywardness in society there are some inherent defect (S1A-006)

(7) …there're some clipping from Sandhyanand (S1A-099)

As Rupp reports, “[o]ne of the noteworthy findings emerging from variationist studies is that alternation between standard concord and nonstandard –s is not erratic or random” (Rupp 2005: 255-256). The variationist work carried out on native Englishes indicated

---

98 See also Hazout (2004) for a very detailed analysis of the syntax of existential constructions.
that, apart from the syntactic considerations mentioned above, there are also some extra-
linguistic factors exerting a strong influence on agreement patterns in different varieties of
English. Previous studies have looked into the potential influence of social factors, such as
gender (Britain and Sudbury 2002, Eisikovits 1991 and Schreier 2002, among others), age
(e.g. Tagliamonte 1998), social class or education (Meechan and Foley 1994, Smallwood 1997 and others) and ethnicity (Britain and Sudbury 2002) on agreement in existential
*there + be* constructions (cf. section 3.2), and situational factors such as the effects of
register characteristics or formality (Crawford 2005). The analysis of the effects of register
and formality as well as the mode of production on subject-verb agreement presented in
chapter 4 supports the findings of Martinez and Martinez (2003) that singular agreement in
existential *there + be* is more frequent in speech than writing and also provides additional
evidence for the result obtained by Smallwood (1997, reported in Walker 2007: 152) that it
seems to be more frequent in less formal speech styles.

The linguistic factors that have been found to condition agreement are subject number,
tense, contractedness, polarity, the syntactic distance of the subject and the verb, as well as
the type of determiner (cf. Walker 2007: 152f., who gives an overview of language-
internal and –external factors featuring prominently in previous investigations of
existential *there + be*).

A special case of non-agreement in existential *there + be* constructions is included as
feature 55 in Kortmann and Szmrecsanyi’s (2004) 76-feature catalogue of morphosyntactic
features in non-standard varieties of Englishes worldwide, namely
“existential/presentational *there’s, there is, there was* with plural subjects.” In their attempt
to identify true candidates for vernacular universals (Chambers, e.g. 2004) they show that
feature 55 “assumes a unique position since it is the only morphosyntactic feature which
has received an ‘A’ classification for all eight British Isles varieties investigated”
(Kortmann and Szmrecsanyi 2004: 1164). In other words, it is “pervasive (possibly
obligatory) or at least very frequent” (Kortmann and Szmrecsanyi 2004: 1142) all over
Britain. In addition, it is rated ‘A’ in eight out of nine American varieties. Most research
into linguistic variation has found that this use of non-standard –s in existential *there + be*
constructions with plural subjects is “one of the most widespread, albeit non-standard,
features of English varieties” (Rupp 2005: 255) and both Chambers (2004) and
Tagliamonte (2006) identify such ‘default singulars’ as the “proto-typical exemplar of a
‘vernacular universal’” (Tagliamonte 2006).
However, almost all of the studies investigating non-standard agreement in existential constructions have put their research focus on native varieties, while considerably fewer studies have been devoted to non-native varieties (e.g. Tagliamonte 2006). With regard to non-native varieties Kortmann and Szmrecsanyi have found that this feature is attested, but neither “among the top features of L2 varieties nor of Pidgins and Creoles” (2004: 1186). Kortmann and Szmrecsanyi (2004) thus do not support the view of Chambers (2004), who identifies ‘default singulars,’ i.e. subject-verb non-agreement with plural subjects, as a candidate for universals of vernacular Englishes but suggest that default singulars are overall a widespread feature of varieties of English which differs with variety type, i.e. feature [55] “is categorical in all L1 varieties” (Kortmann and Szmrecsanyi 2004: 1199), but not in L2 or Pidgin/Creole varieties.

As there is still a lack of investigations into agreement in existential constructions in non-native varieties of English, the present study intends to fill this gap in previous research by systematically analysing non-standard agreement in there + be constructions in five native and non-native varieties of English with a particular focus on the three L2 varieties, Jamaican English, Indian English and Singapore English.

Due to the fact that non-agreement in existential there + be constructions is attested in Englishes all over the world (Kortmann and Szmrecsanyi 2004) it cannot plausibly be explained by innovation. For this reason, this paper will test only the factors substrate influence and vernacular universals or angloversals as potential explanations for variation across the varieties under consideration. The research question that the following two chapters attempt to answer is thus the following. What is the most likely explanation for variation in existential there + be constructions? Are the presence (or absence) and the nature of substrate languages in the respective variety and resulting contact effects the dominant factors causing linguistic variation in existential there + be constructions in Jamaican English (and other ESL countries such as Indian and Singapore English) or can variation be better explained by universal tendencies in varieties of English across the world?

5.3 Data and Methodology

The data basis for the cross-variety comparison of agreement in existential there + be constructions (ETBs) consists of the complete spoken sections of four national components of the International Corpus of English. The variety being focused on is Jamaican English,
and the results from the analysis of ICE-Jamaica are compared to those obtained in ICE-Singapore and ICE-India, representing two other post-colonial contact varieties of English, and to spoken ICE-GB, representing one of the native varieties in question. All these components consist of 300 texts of approximately 2000 words each, which add up to a total of around 600,000 words per variety. The data include a wide variety of text types such as private dialogues (text category S1A), for instance conversations or phone calls, public dialogues (S1B), for example broadcast discussions, unscripted monologues (S2A) such as sports commentaries, and scripted monologues (S2B), amongst them broadcast news. As there is presently no ICE-corpus for American English a different corpus has to be used as the basis for the comparison, which needs to be at least roughly similar to the ICE-corpora in terms of its compilation. The Santa Barbara Corpus of Spoken American English seems to be the best choice, as it has been designed to form part of a future American ICE-component. However, only 166,000 words (83 texts) were available, which were taken predominantly from face-to-face conversations and telephone conversations, and the American data basis is thus considerably smaller than the other corpora. Nevertheless, results obtained from the SBC are used here, but they are compared exclusively to results from the most informal conversations category S1A of the ICE-corpora and not to those obtained from more formal text types, in order to ensure a rough comparability of the results.

Comparative analyses of different types of Englishes are hoped to provide insights into the question whether the pattern in Jamaican English is determined mainly by the presence of the substrate language Jamaican Creole, or whether other factors are more crucial in determining the agreement pattern in educated Jamaican English. If the agreement patterns of all post-colonial varieties under consideration resemble one another closely, substrate influence would seem unlikely as the main motivator, since all varieties have substrates of different types and from different language families. If all three contact varieties pattern similarly but differ considerably from native varieties such as British or American English, then universals of post-colonial Englishes, so-called ‘angloversals’ (Mair 2003: 84) could be a more plausible explanation for the agreement pattern in the Jamaican, Singaporean

99 For a more detailed description of the text categories included in the spoken section of the ICE-corpora consult Appendix 1 or the homepage of the International Corpus of English project at <http://www.ucl.ac.uk/english-usage/ice/>.

100 For more information on the Santa Barbara Corpus of Spoken American English consult their website at <http://www.linguistics.ucsb.edu/research/sbcorpus.html>.

101 Section 5.3.1 gives a short description of the linguistic situation and the predominant substrate languages in the varieties chosen for this comparative study.
and Indian sub-corpora. The current chapter will focus particularly on existential \( \text{there} + \text{be} \) constructions, comparing the agreement patterns in the L1 and L2 Englishes, and in chapter 6 the perspective will be shifted to an investigation of the various means of expressing existence in the substrate languages, in order to explore the possibility of substrate influence on existential constructions in the respective varieties of English.

### 5.3.1 Data – Variety-Choice for the Study of Inter-Corpus Variation

For the comparative study of existential constructions, two post-colonial varieties from the International Corpus of English were chosen which were historically British-oriented, and are thus well suited for a comparison with Jamaican English, and which have as wide a range of substrate languages as possible. The next section provides a brief overview of the history of Singapore English and Indian English and describes the current linguistic situation of these countries, drawing particular attention to the background languages present in the respective region, in order to justify why these Englishes are here considered to be the most suitable ones for this study.

#### Singapore English

Until the 19th century Singapore was part of the Sultanate of Johore. The year 1819, when Sir Stamford Raffles signed a contract to obtain Singapore as a trading post for the British East India Company, tends to be considered as a milestone in Singapore history (e.g. Lim and Foley 2004: 2). Singapore’s location near the trade routes of Southeast Asia and China was very attractive for settlement, so that immigrants from Penang and Malacca, China, Indonesia, Sri Lanka and India, as well as some Arabs and Eurasians let the population numbers increase rapidly (cf. Sand 2005: 38). In 1826 Singapore became part of the newly formed Straits Settlements consisting at the time of Penang and Malacca. After having been governed through India for a considerable period (Gupta 1998: 106f.), it became a British Crown Colony in 1867 and remained British until the Japanese defeated the British army in 1942 and occupied the country for three years. In 1945 Singapore became British again and gained self-government within the British Empire in 1959 (cf. Sand 2005: 39). After having been part of the Federation of Malaya from 1963 onwards Singapore became independent in 1965.
Due to the presence of the *lingua franca* Bazaar Malay, which was spoken by locals and colonial agents of the East India Company alike (Ansaldo 2004: 141), there was no need for an English-based pidgin in Singapore. However, many of the Eurasians coming to Singapore had been in contact with the British Empire, and thus with English, so that some form of English was spoken in Singapore even before the first English schools were established in Penang in 1787 (Khoo 1996: 113). A large proportion of children being exposed to English-medium education until the 19th century were Eurasians (see Gupta 1994), while the non-European inhabitants tended to speak very little or no English at all (Sand 2005: 39). Those educated in English belonged to an elite which was socially more mobile and able to span a much wider occupational range than the remainder of the population (cf. Lim and Foley 2004: 3). Before the Japanese occupation, English was taught exclusively in English-medium schools, but by the 1920s and 30s many of the Non-English-medium schools were also teaching English as a subject (see Lim and Foley 2004: 4). It was not until the 1950s, however, that education in English became the norm in Singapore and after independence English was established as a compulsory language in schools (Lim and Foley 2004: 4) for all children.

The current linguistic composition of Singapore is very complex, with English being just one out of a large number of available linguistic codes regularly used by the population. As Table 15 shows, 76.8% of the roughly 3.3 million residents in Singapore are of Chinese ethnicity, 13.9% are Malays, 7.9% are Indian and 1.4% are people of other races, including Eurasians, Europeans and Arabs (McArthur 2002: 338).

<table>
<thead>
<tr>
<th>Year: 2000</th>
<th>Chinese</th>
<th>Malays</th>
<th>Indians</th>
<th>Others</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographic Characteristics</strong></td>
<td><strong>Singapore Residents (*'000)</strong></td>
<td>2,505.4</td>
<td>453.6</td>
<td>257.8</td>
<td>46.4</td>
</tr>
<tr>
<td><strong>Ethnic Composition (%)</strong></td>
<td>76.8</td>
<td>13.9</td>
<td>7.9</td>
<td>1.4</td>
<td><strong>100</strong></td>
</tr>
<tr>
<td><strong>Language most frequently spoken at home (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>23.9</td>
<td>7.9</td>
<td>35.6</td>
<td>68.5</td>
<td><strong>23.0</strong></td>
</tr>
<tr>
<td>Mandarin</td>
<td>45.1</td>
<td>0.1</td>
<td>0.1</td>
<td>4.4</td>
<td><strong>35.0</strong></td>
</tr>
<tr>
<td>Chinese Dialects</td>
<td>30.7</td>
<td>0.1</td>
<td>0.1</td>
<td>3.2</td>
<td><strong>23.8</strong></td>
</tr>
<tr>
<td>Malay</td>
<td>0.2</td>
<td>91.6</td>
<td>11.6</td>
<td>15.6</td>
<td><strong>14.1</strong></td>
</tr>
<tr>
<td>Tamil</td>
<td>--</td>
<td>0.1</td>
<td>42.9</td>
<td>0.2</td>
<td><strong>3.2</strong></td>
</tr>
<tr>
<td>Others</td>
<td>0.1</td>
<td>0.3</td>
<td>9.7</td>
<td>8.2</td>
<td><strong>0.9</strong></td>
</tr>
</tbody>
</table>

Adapted from Geok 2001: viii-ix.102

102 For the original table comparing demographic characteristics and the linguistic make-up of the Singaporean population in the Census 2000 with the 1990 census see Appendix 2 or consult the website of the Singapore Department of Statistics at www.singstat.gov.sg.
The languages most frequently spoken at home are Mandarin, which is regularly used by 35% of the total population, other Chinese dialects such as Hokkien, Teochew or Cantonese, spoken by almost 24% of the population, as well as English, which is the preferred home language of 23% of the Singaporeans, but particularly of those Singaporeans who are not of Chinese, Malay or Indian descent. Further languages spoken in the country are Tamil, and other Dravidian languages, Malay, the home language of 14.1% of the population, and Bazaar Malay (Bahasa Pasar), a Malay-based pidgin which has always served as the main lingua franca in the region (see Lim and Foley 2004: 3). However, only the national language Malay, and Chinese, Tamil and English have the status of official language in Singapore. The official languages are typologically rather different, in that English is a rather analytic language with little inflectional morphology which belongs to the Germanic branch of the Indo-European language family, Mandarin and other Chinese dialects are also isolating (or analytic) languages, but belong to the Sinitic language family, and Malay is an agglutinative, i.e. synthetic, Austronesian language. They all typically have Subject Verb Object (SVO) word order. By contrast, Tamil is an agglutinative Dravidian language having the typical word order SOV.

In contrast to the other official languages, English is used by all sections of the population, as it is the language of law, commerce and administration and the main medium of education\(^{103}\) in universities as well as in schools, but it can also be used for informal purposes in private domains. However, the variety of English chosen is not the same for every communicative purpose. The current situation of English in Singapore is similar in this respect to that in Jamaica, as described in section 2.1.1, where several varieties are available whose choice depends on various social and situational factors.

It is commonly assumed that the English language in Singapore comes in two distinct shapes: the Standard English primarily used in the public domain (media, administration, etc.) and the vernacular Colloquial Singapore English\(^{104}\) (CSE). The relation of these two major varieties is, however, not clear. Two models have been put forward to explain the complex interaction of Standard Singapore English (SSE) and CSE. One approach, taken for example by Gupta (1994, 2001) considers SSE the H(igh) and CSE the L(ow) variety

---

\(^{103}\) See Sand (2005) for a description of the language policy implemented in the school system, where “each child is assigned an official ‘mother tongue’ based on ethnicity, which means that Chinese children will study Mandarin as a second language, Malay children study Malay and Indian children study Tamil. The ‘mother tongue’ is not necessarily the child’s L1” (p. 40).

\(^{104}\) Colloquial Singapore English is often referred to as ‘Singlish’ in the literature. As this term tends to have negative, derogatory connotations of being ‘non-standard’ or ‘ungrammatical’ (cf. Lim and Foley: 2004, footnote 1), it will be avoided in the present paper.
in a case of classic diglossia (cf. Ferguson 1959), which are functionally in complementary
distribution, with H being used exclusively in public domains such as politics and print and
L being used for informal private purposes only. However, Ansaldo rejects diglossia as a
model for the sociolinguistic status of Singapore English, as it seems a “rather simplistic
and somewhat inadequate (outdated) description […]” (2004: 145). Instead, he proposes
that Singapore English comprises a continuum of lects\(^{105}\) ranging from an acrolectal level
that is very close to Standard English, to a basilectal level characterized by “almost
demonstrates, however, that both approaches have considerable weaknesses in accounting
for the complex interaction of sub-varieties of Singapore English and proposes a combined
model that better captures the variation found in his data by localising the individual
speakers on a scale of proficiency in the High variety. Leimgruber proposes “two diglossic
varieties, which differ according to the speaker’s level of H proficiency, and which are
located on a linguistic continuum from SSE to CSE” (Leimgruber 2008).
What is of particular interest here, however, is Ansaldo’s finding of large numbers of
hybrid features, i.e. features “derived from the interference of typologically different
linguistic systems and universal tendencies of shift that can be observed in contact
situations” (2004: 127) in his data. He manages to trace two main sources of substrate
influence on the development and current shape of Singapore English. One such source is
the lingua franca used throughout much of Singaporean history, namely Bazaar Malay, and
not the standard variety High Malay (Ansaldo 2004: 131). He locates the second source of
linguistic influence in Chinese varieties such as Hokkien, a Min language which was
already spoken in Singapore in the 15\(^{th}\) century, long before the first European colonists
arrived, and Cantonese, a Yue language which entered the country more recently but is
“extremely vital in Singapore to this date” (2004: 132). Ansaldo concludes that CSE can be
described as “having an English lexifier base and an otherwise isolating typology with
mainly Austronesian (Malay) and Sinitic (Hokkien and Cantonese) substrate influences”

\(^{105}\) For a detailed discussion of the post-creole continuum, i.e. a continuum of sub-varieties of English
ranging from an acrolect to a basilect with several mesolectal varieties in between, see DeCamp (1971) and
Platt (1975) who uses this continuum model to account for the sociolinguistic environment of Singapore.
Indian English

The second post-colonial contact variety chosen for comparison with Jamaican English is Indian English, which was, like Jamaican and Singapore English, historically British-oriented. Although an ambassador of Alfred the Great is assumed to have been the first speaker of English to visit India, as stated in 884 in the Anglo-Saxon Chronicle (Kachru 1994: 501), the language entered the country only much later, after Elizabeth I. had granted the trade monopoly with East India to a number of merchants from London in 1600, which resulted in the foundation of the East India Company (Hickey 2004: 537). While Britain was not the only European power involved in trade relations with India, the other three being France, Portugal and the Netherlands, the British East India Company came to be the most important of the various East India Companies after establishing so-called ‘factories,’ i.e. trading stations, in Surat (1612), Madras (1639-40), Bombay (1674) and Calcutta (1690) (Hickey 2004: 539). According to Hickey (2004: 539), the following three factors were particularly important for the further spread of English (cf. Kachru 1983), among the native population in India:

1) The activity of Christian missionaries in the seventeenth century, who aimed at but did not succeed in christianizing large sections of the population, but who also established the first English-medium schools, for instance, St Mary’s Charity Schools in Madras in 1715 and Bombay in 1719 (Hickey 2004: 539);

2) The local ‘demand’ for English use by higher sections of the society. After the British came to India they looked for native Indians to assist them in administering the country and introduced a policy to establish an elite of high-caste Indians who should be “Indian in blood and colour but English in taste, in opinions and morals and intellect” (Thomas B. Macaulay, quoted in Mehrotra 1998: 4). As Indians with a good command of English were seen as the social elite, this policy resulted in positive attitudes toward English as the language of upward social mobility which opened the doors to knowledge of the West (Hickey 2004: 541). Consequently, socially ambitious Indians preferred higher education schools in which English was the medium of instruction and English was thus given preference over Sanskrit, Arabic, Persian and indigenous languages (Hickey 2004: 541). An example is the Bengali Hindu social reformer Raja Rammohan Roy, who expresses his strong wish for the British government to employ [...] European gentlemen of talents and education to instruct the natives of India in mathematics, natural philosophy, chemistry, anatomy,
and other useful sciences, which the natives of Europe have carried to a degree of perfection that has raised them above the inhabitants of other parts of the world. (quoted in Kachru, 1985: 34)

In the controversy over the appropriateness of different educational systems for India, Roy represents a group called the ‘Anglicists’ or ‘Occidentalists,’ who favoured English as the medium of education in Indian schools, while another group, the ‘Orientalists’ favoured local languages such as Sanskrit as the medium of instruction (Hickey 2004: 540). Thomas Babington Macaulay, a member of the Supreme Council of India, decided the language matter in favour of English in an official Minute in 1835 (Sand 2005: 27).

3) This decision to make English the medium of education was the third main factor in the spread of English in India. The first English-medium universities were founded in Bombay, Calcutta and Madras in 1857, and by 1882 more than 60% of all Indian primary schools used English as the medium of instruction (Kachru 1994: 507).

During “the British raj ‘sovereignty’” (Hickey 2004: 542), which lasted from 1765 until independence in 1947 (Raychaudhuri 1996), English also became the primary language in administration and education on all levels (Gupta 1996). The role that English was to play in sovereign India was hotly disputed, and this is clearly reflected in Article 343 of the 1950s constitution, which aimed at finally replacing English with Hindi in its official functions as the language of education, administration and legislation by stating that “[t]he English language shall continue to be used for a period of 15 years from the commencement of this Constitution for all official purposes of the Union for which it was being used immediately before such commencement” (quoted in Mehrotra 1998: 6). However, the Official Language Commission of the Government of India explicitly declared that it did not recommend

that any restriction should be imposed for the present on the use of the English language for any of the purposes of the Union… It is not suggested that English be rejected merely because it is a foreign language for we entirely agree that a language is not the property of any particular nation, and obviously it belongs to all who can speak it. (quoted in Mehrotra 1998: 6)

It was thus decided not to replace English by Hindi in all functions. In fact, when, in the 1960s, a three-language-formula was introduced in the education system in order to regulate the use of languages in the education system (cf. Sand 2005: 29), English was institutionalized as one of three languages to be taught to all children and thus the role of
English in the education system was strengthened. Students all over India now learn English and the national language Hindi, as well as one regional language, normally the official state language. According to Sand, “it was originally hoped that the students from Hindi-speaking areas would learn an additional language, but in effect the formula is reduced to two languages in those areas” (2005: 29). After much political pressure from non-Hindi-speaking Indians, the Official Language Amendment Bill in 1967 granted English the status of an “additional or associate official language” (Sand 2005: 28).106

Even nowadays India has a very interesting and extremely complex socio-linguistic make-up, which is, however, totally different from that in Singapore or Jamaica and thus well-suited for the present comparative study. As mentioned above, India has two official languages, English and Hindi. Hindi is a highly synthetic language that belongs to the Indo-Aryan branch of the Indo-European language family. Its normal word order is SOV, but it can very flexibly be changed for emphasis. Hindi is the national language of India which has the highest number of official functions, namely in central administration, the armed forces, the media, etc., and it functions as the official language of seven states of the Union (Hansen et al. 1996: 216f.). English, on the other hand, is the second official language, and the state language in five states of the Union and eight Union territories (Sand 2005: 29). Moreover, it is the language of parliament, trade and commerce, administration, higher education and inter-state communication (cf. Mehrotra 1998).

According to McArthur (2002: 312), about 100-200 million people in India were regular speakers of English in 2002, but it is spoken as a mother tongue only by a very small fraction of the population. In the 2001 census, the number of mother tongue English speakers is given as 226,449 (0.02% of the population) (Census of India 2007). This means that English is only regularly used by 5-10 percent of the total population and is thus “still a language only spoken by a small elite” (Sand 2005: 30).

The multilingual environment in India where so many structurally different languages are in contact naturally leads to a high amount of linguistic variation, not only caused by regional differences in the sociolinguistic make-up, but, as Sand (2005: 30) describes, there is also a cline of proficiency in English, which ranges from a low variety Pidgin English (Mehrotra 1998: 5) over the middle varieties Butler English (cf. Hosali 2000) and Boxwalla English (cf. Kachru 1994: 512f.), which can be described as restricted codes, to

---

106 Hickey (2004) describes the spread of English and particularly the relevant political developments in more detail. See also Mehrotra (1998) for a sketch of the history of English in India, with particular focus on the role of English in Indian media and literature.
the high variety Babu English (cf. Kachru 1994: 509-511). In addition there is formal Standard English. According to Mehrotra, educated Indians’ speech is characterized by a mixture of English with Indian elements, particularly in informal situations, resulting in hybridized forms of English which are often referred to as *Hinglish, Tamlish* or *Indish* (Mehrotra 1998: 14).

According to the Census of India\(^{107}\), the total population of India in 2001 was 1.028 billion. The number of Indians speaking Hindi as their mother tongue\(^{108}\) is given as 422 million, or 41 percent of the population. In the census of 1991, 18 languages were officially recognized “scheduled languages” but with the 100\(^{th}\) Amendment to the Constitution of India, four languages were added.\(^{109}\) About 96 percent of the inhabitants have one of the 22 scheduled languages as their mother tongue, the remainder being included in the census as ‘other languages’ (Census of India 2007). Table 16 shows the five languages with the highest number of mother tongue speakers.

<table>
<thead>
<tr>
<th>No.</th>
<th>Language</th>
<th>Mother tongue speakers</th>
<th>% of total population</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hindi</td>
<td>422,048,642</td>
<td>41,0%</td>
</tr>
<tr>
<td>2</td>
<td>Bengali</td>
<td>83,369,769</td>
<td>8,1%</td>
</tr>
<tr>
<td>3</td>
<td>Telugu</td>
<td>74,002,856</td>
<td>7,2%</td>
</tr>
<tr>
<td>4</td>
<td>Marathi</td>
<td>71,936,894</td>
<td>7,0%</td>
</tr>
<tr>
<td>5</td>
<td>Tamil</td>
<td>60,793,814</td>
<td>5,9%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>712,151,975</td>
<td>69,2%</td>
</tr>
</tbody>
</table>

(adapted from Census of India 2007)\(^{110}\)

Table 16: The five scheduled languages with the highest numbers of mother tongue speakers in the 2001 census of India

---

\(^{107}\) For more information consult the website of the Census of India at [http://www.censusindia.gov.in/](http://www.censusindia.gov.in/).

\(^{108}\) The mother tongue is defined as “the language spoken in childhood by the person’s mother to the person” or, in case of an early death of the mother, “the language mainly spoken in the person’s home in childhood” (Census of India 2007, General Note, [http://www.censusindia.gov.in/Census_Data_2001/Census_Data_Online/Language/gen_note.htm](http://www.censusindia.gov.in/Census_Data_2001/Census_Data_Online/Language/gen_note.htm)). The exact number of mother tongues in India is hard to determine and numbers presented in the literature differ considerably. While the 1961 Census of India speaks of 1652 claimed mother tongues from four main language families, Indo-European, Dravidian, Austro-Asiatic and Sino-Tibetan, the SIL Ethnologue database (Gordon 2005) gives a much smaller number and states that of the 428 languages listed for India, 415 are living languages and 13 are extinct. The discrepancy between these numbers may have been caused by the fact that the census number denotes not necessarily separate languages or dialects but all terms used by the individuals to label their speech, while the Ethnologue database lists only languages and/or dialects.

\(^{109}\) The scheduled languages in alphabetical order (including the four languages added in 2003 which are marked in italics) are Assamese, Bengali, *Bodo, Dogri, Gujarati, Hindi, Kannada, Kashmiri, Konkani, Maithili, Malayalam, Manipuri, Marathi, Nepali, Oriya, Punjabi, Sanskrit, Santali, Sindhi, Tamil, Telugu and Urdu* (Census of India 2007).

\(^{110}\) For the full table on the 22 scheduled languages in India ordered by the number of mother tongue speakers see also Appendix 3. A more extensive list of languages (122) and the mother tongues (234) with speakers’ strength of 10,000 and above at the all India level is given by the Census of India online at [http://www.censusindia.gov/Census_Data_2001/Census_Data_Online/Language/Statement1.htm](http://www.censusindia.gov/Census_Data_2001/Census_Data_Online/Language/Statement1.htm).
Hindi and Bengali and Marathi, three of the five the languages with most mother tongue
speakers, belong to the Indo-Aryan branch of the Indo-European family which is more
prevalent in the northern part of India, while Telugu and Tamil are Dravidian languages
(Census of India 2007) more common in the South.

As neither Singapore nor India has a creole substrate, a comparative investigation of
agreement patterns in existential *there + Be* constructions in these varieties can be a
valuable tool in exploring regularities in the development of post-colonial contact
Englishes. If all varieties of English considered here resemble one another, then substrate
influence cannot have caused the patterns, and tendencies applying universally in post-
colonial Englishes might more plausibly explain the similarities. If, however, there are
crucial differences between the varieties that are clearly traceable to the background
languages, then the patterns may be due to substrate influence. The variety being focused
on is still Jamaican English. The question is whether it is indeed strongly influenced by
Jamaican Creole\(^{111}\), as postulated by Mair (2002: 32), or whether the patterns in the
Jamaican data are characteristic of post-colonial Englishes generally irrespective of the
background languages and the type of contact history present in the respective regions.

**British and American English**

British and American English are the two native reference varieties used for comparison in
the present study. Of all native varieties of English these are most relevant for an
investigation of Jamaican English. British English played an important role in Jamaica for
a long period of time as the historical norm imposed by the British colonial administration
and it was used as the target variety in the education system. In fact, until quite recently,
standard Jamaican English was not regarded as a variety in its own right but was
considered to be identical to standard British English. After independence, however,
influence from British English has been decreasing, and some teachers and applied
linguists (e.g. Christie 1989) and later sociolinguists (e.g. Mair 2002) found Jamaican
English to be moving away from the postulated British norm (see Sand 1999: 13-14).
American English has been presumed to exert an influence on Jamaican English due to its
geographical proximity and the economic and cultural attractiveness of a global power. In
particular, American English is THE language of technology, the internet, and the media.

\(^{111}\) Jamaican Creole is an analytic SVO language of the creole type with virtually no inflectional morphology
(McWhorther 2005) and thus differs from all other background languages considered here.
In the past, however, contact between Jamaicans and speakers of American English was less direct than that with speakers of British English. While Jamaicans were directly exposed to the English spoken by the British mother tongue speakers until independence from Britain, American English was used mainly in the media. While Jamaicans had a long history of sporadic contact with speakers of American English especially since the construction of the Panama Canal in the late 19th century, direct contact has strengthened considerably only in recent decades. Therefore, the ratio of influence of these varieties on Jamaican English has been changing, as British English influence has been steadily decreasing after political independence while American English influence has been greatly increasing.

The comparison of Jamaican English with both native varieties will show which of these Englishes exerts a stronger influence on agreement patterns in existential *there + be* constructions (ETBs) and how it does so. Substrate influence can only be a plausible explanation for non-agreement in existential *there + be* constructions of these non-native varieties if the patterns in these L2s differ markedly not only from both native varieties but also from one another112, as different substrates are unlikely to lead to similar influences on the English spoken in the respective regions (cf. Mesthrie 2004).

Before turning to the methodology adopted in this investigation, it is important to note that the compilation of the ICE-corpora tends to be slightly biased toward formal language use. This was found for British English (see Martinez and Martinez 2003: 268 and footnote 11) in their comparison of results from ICE-GB with the BNC, but it is also assumed to be true for the non-native varieties, the reason being that English tends to be used for official and formal functions in Jamaica, India and Singapore and thus constitutes the more formal language choice in these varieties. The American data from the SBC is therefore expected to show higher overall non-agreement rates than the data from the ICE-corpora.

The next section provides a short overview of the methods applied in the analysis of existential *there + be* constructions.

---

112 Substrate influence is here analysed quantitatively. It is assumed that if substrate equivalents of such different structures as the ones described here do exert an influence on agreement in existential *there + be* constructions, then this should be reflected in the frequencies with which non-agreement is found in the respective categories. The exact nature of such influence will be examined in chapter 6.
5.3.2 Methodology

For the investigation of existential constructions in the five native and non-native varieties under consideration, a WordSmith search of THERE was carried out in the spoken sections of the respective ICE-corpora and all instances of there + be were extracted. Excluded from analysis were all questions, e.g. (8), and tag questions, sentences with modal auxiliaries (e.g. there must be…), semi-auxiliaries (have to),\(^{113}\) or instances where there had a locative or deictic function (2), but also unclear cases, e.g. structurally incomplete text units caused by hesitation or repairs in which there was either no post-verbal noun phrase present (9) or it was more likely the subject of another clause (10).

(8) **Was there** enslavements (ICE-JA, S1B-001)

(9) Yeah but then **there is** also you know people think that… (ICE-JA, S1A-008)

(10) …among this group uh **there are** uh the group represents about… (ICE-JA, S2B-048)

Excluded were also cases in which a post-verbal adverbial there was used in place of dummy there, as in (11). Such cases occurred exclusively in Indian English.

(11) Various companies **are there** and… (ICE-IND, S1B-055)

Text units in which the missing post-verbal noun phrase was disambiguated by the context were however included (12).

(12) Is there a parent teachers association. Yes **there is** (ICE-JA, S1A-096)

The instances of there + be that were included in the analysis were coded as follows. For each token I noted the grammatical number of the subject and verb and whether the verb agreed in number with its subject, as in (1) or not, as in examples (3) – (7).

Moreover, I coded the tense of the verb, i.e. whether the verb had present or past reference, and polarity, i.e. whether then sentence was positive or negative. For the analysis of contractedness, only positive sentences were considered, in order to facilitate investigations into agreement differences between the full forms there is/are + noun versus

\(^{113}\) The present analysis focuses on existential constructions with a form of be. Non-agreement with have is discussed in the investigation of subject-verb agreement with lexical verbs in sections 4.3.1 and 4.4.1. Note, however, that none of the instances of non-agreement with have was an existential construction such as there have to be + noun phrase or there have been + noun phrase. Existential there constructions with other verbs, such as appear or seem are also excluded from analysis. Cf. footnote 97. Modal idioms such as be to are not attested in the spoken corpus sections under consideration.
the contracted forms *there’s*/*there’re* + noun. In other words, negative forms such as *there isn’t*/*aren’t*/*wasn’t* + noun were not counted as instances of contracted *there + be*.

As these four grammatical features, subject number, tense, polarity and contractedness, were identified as potentially influential in determining agreement and non-agreement in existential *there + be* constructions, these will be the focus of the present study. The analysis yielded the following results.

### 5.4 Agreement in ETBs in Native and Non-Native Varieties of English

#### 5.4.1 Agreement in the Varieties

Of the 2052 instances of existential *there + be* constructions in the British corpus, 132 (6.4%) had to be excluded from analysis as unclear cases. In the American data, 26 out of a total of 406 instances (6.4%) were not analysed, in the Jamaican sub-corpus 33 out of 1430 tokens (2.3%) were excluded, in the Indian data there were 1843 instances of existential *there + be*, 92 (5.0%) of which had to be disregarded, and, last but not least, 53 out of 1596 tokens (3.3%) in Singapore English were not taken into account. The total number of token included in the analysis, as well as the distribution of agreement and non-agreement, is presented in Table 17 for the five varieties under consideration.

<table>
<thead>
<tr>
<th>Variety of English</th>
<th>Agree</th>
<th>Non-agree</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tokens</td>
<td>%</td>
<td>Tokens</td>
</tr>
<tr>
<td>British English</td>
<td>1786</td>
<td>93,0%</td>
<td>134</td>
</tr>
<tr>
<td>American English</td>
<td>314</td>
<td>82,6%</td>
<td>66</td>
</tr>
<tr>
<td>Jamaican English</td>
<td>1328</td>
<td>95,1%</td>
<td>69</td>
</tr>
<tr>
<td>Indian English</td>
<td>1651</td>
<td>94,3%</td>
<td>100</td>
</tr>
<tr>
<td>Singapore English</td>
<td>1436</td>
<td>93,1%</td>
<td>107</td>
</tr>
</tbody>
</table>

Table 17: Distribution of agreement and non-agreement in existential *there + be* constructions

The number of occurrences of existential *there + be* constructions is highest in the British corpus, with 1920 tokens, and lowest in the American data, with only 380 tokens, with the other varieties ranging in between. This extreme difference is due to the data basis used for
the analysis, as the Santa Barbara Corpus of Spoken American English (SBC) not only comprises exclusively conversational data, but it is also considerably smaller, consisting of only 83 texts as compared to 300 in the sub-corpora of the other four varieties. The American data will therefore only be used for comparison of results relating to the conversational data. American English displays a rate of non-agreement that is much higher than in the other varieties (17.4%).\textsuperscript{114} The overall rates of non-agreement in existential \textit{there + be} constructions in the varieties under consideration, excluding American English, range between 4.9% in Jamaican English and 7.0% in British English. There is no clear difference in non-agreement rates between the types of varieties. Although the native varieties display the highest non-agreement rates (7.0% in BritE and 17.4% in AmE), the difference to Singapore English, the variety with the highest non-agreement rate (6.9%) among the L2s, is negligible. All in all, non-agreement in existential sentences cannot be considered a frequent phenomenon, at least not across the whole spoken corpus sections.

However, the investigation of register variation presented in chapter four provided evidence for the assumption that non-agreement patterns are strongly influenced by situational factors and thus differ across registers, being more frequent in informal than formal situations. This general tendency can also be assumed to appear in agreement patterns in existential constructions if situational factors are taken into consideration.

### 5.4.2 Situational Factors

Figure 19 shows the overall rates of agreement and non-agreement in \textit{there + be} constructions in the respective varieties, once across all text categories\textsuperscript{115} and once only in the conversational category S1A.

\textsuperscript{114} This difference may be due to the different make-up of the data. It is, however, not clear to what extend the compilation of the corpus might be responsible for the high rates of non-agreement in American English in comparison with the other corpora. In view of this uncertainty concerning the comparability of the data, generalizations will have to be made very carefully.

\textsuperscript{115} Except for American English, where the results are based only on conversational data.
As mentioned above, the differences between the varieties are rather small when all text categories, i.e. the average rates of non-agreement in existential *there + be* constructions, are considered. If only the conversational category S1A is analysed, the non-agreement rates increase considerably from 7.0% to 12.4% in British English and from 6.9% to 10.7% in Singapore English, while remaining fairly stable in Jamaican English (4.9% in the whole data set versus 4.85% in S1A) and Indian English (5.7% in the whole sub-corpus as compared to 6.0% in conversations). It appears that the rates in Indian and Jamaican English are different from those in the native varieties and this might be an indication of non-agreement in *there + be* constructions being in some way related to the type of variety. However, Singapore English resembles the native varieties more closely than the L2s. In order to explain this pattern, a more detailed analysis of non-agreement in existential constructions is necessary. Turning to the rates of non-agreement across four different text categories, Figure 20 shows that the amount of variation in non-agreement is highest in category S1A.
In all varieties, non-agreement is higher in the most informal category S1A, i.e. private dialogues, than in the highly formal text type S2B, i.e. scripted public monologues such as broadcast news. In the latter text type, non-agreement is least frequent. This tendency towards register variation appears in all varieties, but otherwise no general trends can be observed for all varieties. It is only British English that shows the expected pattern in that non-agreement rates decrease steadily with increasing relative level of formality, being highest in the private dialogues (S1A), considerably lower in public dialogues (S1B) and unscripted monologues (S2A), and lowest in the category of scripted speeches (S2B). It has to be noted, however, that the text categories are rather heterogeneous, which means that the text categories of medium formality resemble one another rather closely in various respects and this may explain why no clear differences between them are visible with regard to non-agreement rates. The only general statement that can be made is that variation is highest in the most informal of the text categories, and that it tends to decrease, as it is highest in the informal conversational data and lowest in the highly formal speeches.

A statistical Chi square test for text category and non-agreement in *there* + *be* constructions yields $\chi^2 = 28.879$, ($p = 0.000$), and this result shows that there is indeed a highly significant correlation between text type and the amount of non-agreement found in existential *there* + *be* constructions.

---

116 American English is not included in this chart, as only conversational data was available for this variety.
What is surprising is that it is the native varieties British and American English (17.4% non-agreement in S1A) and not the non-native varieties that show the highest rates of non-agreement in S1A, although the latter are assumed to be subject to substrate influence particularly in informal language. There are thus no traces of direct contact influence on the amount of non-agreement in this most informal text category, which has previously been identified as most susceptible to substrate influence (cf. Jantos, forthc.). A thorough investigation of the grammatical factors subject number, tense, contractedness and polarity will provide a clearer picture of (non-) agreement patterns in existential *there + be* constructions.

5.4.3 Subject Number

Figure 21 illustrates the effect of the grammatical number of the post-verbal noun phrase, i.e. the notional subject, on agreement in existential *there + be* constructions. It indicates that non-agreement is generally rather infrequent with rates between one and around 20%, except in American English where 52% of all verbs with plural subjects display non-standard agreement.

The data show that non-agreement with plural subjects is consistently higher than with singular subjects in all varieties under consideration, and this result supports previous research findings reported above (cf. section 5.2). Examples (13a) and (b) represent cases
of agreement and non-agreement with singular notional subjects, respectively, while corresponding examples with plural subjects are presented as (14a) and (b).

(13a) I know for a fact that **there is energy** (ICE-JA)

(13b) …**there are** so much emphasis on children (ICE-Sing)

(14a) **There are** lots of **windows** (ICE-GB)

(14b) Ah **there is** lots of **seats** (ICE-Sing)

There is a clear effect of subject number on non-agreement rates, and this is confirmed by a highly significant Chi square test over all varieties ($\chi^2 = 22.735$, df=4, p=0.000). Non-agreement with singular subjects is completely absent from the American data and ranges between 1%\(^{117}\) (BrE) and 5.1% (IndE) in the other varieties. By contrast, with 7%, Indian English is the variety that has the lowest rate of non-agreement with plural subjects. Jamaican English, Singapore English and British English take middle positions with 11.6%, 14.3% and 20.2% non-agreement, respectively, and in American English a singular verb form is used with 52% of the plural subjects in *there* constructions. The effect of subject number is thus more pronounced in the native varieties than in the L2 varieties. Chi square tests for each variety individually yield highly significant results (p = 0.000) for all varieties except Indian English, where the result is slightly below the critical value of 2.706 at p = 0.1.\(^{118}\)

These relatively low rates show that non-agreement in existential *there + be* constructions is neither categorical nor even the preferred form used by the majority of speakers in any of the varieties.\(^{119}\) However, it is available as a non-standard option to speakers of all Englishes considered. The results described above confirm the conclusion of Rupp (2005), Chambers (2004) and Tagliamonte (2006) that default singulars are a pervasive feature that is attested in Englishes around the world, but this feature is more pervasive in L1 varieties than L2s in my data, thus supporting Kortmann and Szmyrecsanyi’s (2004) results. While the findings indicate that non-agreement in existential constructions is not the norm in any

---

117 This percentage means that of all verbs occurring in existential *there + be* constructions with singular notional subjects 1% have a plural form.

118 The Chi square results for subject number in the varieties (df = 1 in each case) are as follows. British English: $\chi^2 = 235.0$, American English: $\chi^2 = 158.6$, Jamaican English: 63.2, Indian English: 2.6 and Singapore English: 71.7. The critical value at p = 0.005 is 7.879.

119 Except in American English, where more than half of the *there* constructions with plural post-verbal subjects in my data are instances of non-agreement. However, the different data compilation, i.e. the much smaller size of the corpus and its restriction to conversational domains, makes it impossible to compare this result to those obtained for the other varieties. I will therefore refrain from drawing conclusions about existential *there + be* constructions in American English, as they would inevitably be premature and overgeneralized.
of these varieties, which contrasts with Meechan and Foley’s (1994) results for Canadian English, the tendencies are stable and my analysis thus supports Martinez and Martinez (2003). The next grammatical factor that will be analysed is tense.

### 5.4.4 Tense

Tense has been shown in previous research to influence agreement rates in existential *there* + *be* constructions (cf. the summary by Walker 2007: 152f.). Figure 22 indicates that there is indeed an influence of tense on agreement in the varieties under investigation.

![Figure 22: Non-agreement by tense](image)

In all varieties except American English, non-agreement occurs in less than 10% of all verbs irrespective of tense. The high rate of non-agreement in the American data, 19.2%\(^{120}\) with present tense verbs and 12.6% in past tense, may be a result of the corpus being compiled differently from the other corpora, but it may also reflect a tendency toward more non-agreement in *there* constructions in American English. A more detailed analysis of existential *there* + *be* constructions in American English as compared to other varieties would have to be carried out on the basis of comparable corpora in order to provide a clearer picture of this apparent tendency.

What is more interesting than the overall rates of non-agreement is the difference in rates between both tenses. In the native varieties, non-agreement is higher when the verb is in

---

\(^{120}\) This percentage represents the amount of non-agreement in all present tense verbs in the American data.
present tense (19.2% in AmE and 8.1% in BritE, see (15)) rather than past tense (12.6% in AmE and 2.9% in BritE, see example (5), which is reproduced below).

(15) yes there are one topic uhm in every paper (S1A-088)

(5’) I think there was four books (S1A-064)

In the three non-native varieties, this is not the case. Instead, differences between the tenses are extremely small. In SingE and IndE, the amount of non-agreement is slightly lower in present tense verbs (6.5% in SingE and 5.6% in IndE) than in past tense (8.0% in SingE and 6.4% in IndE). In Jamaican English that difference is negligible, with non-agreement being around 5% in both present and past tense.

Chi square tests are, however, insignificant for all varieties except British English, where it is highly significant at the 0.000 level. Surprisingly, even the test for American English is insignificant, although the percentages seem to point to a clear preference of non-agreement in sentences with present tense reference. The result obtained here for American English will have to be tested on the basis of a larger corpus, in order to see whether or not tense determines the rate of non-agreement in existential there + be constructions in this variety.

Hay and Schreier’s (2004) finding that there is a strong preference toward non-agreement in verbs with past tense reference of be cannot be confirmed on the basis of my data for the varieties considered here. In contrast to Deuber (2009a), who observed the infrequency of non-agreement in 40 conversational texts from ICE-Jamaica particularly in the past tense, I found that non-agreement in the whole spoken section of the same corpus, which served as the basis for the present study, turns out to be equally infrequent in present and past tense.

The results for tense suggest that the patterns found in the data might be somehow related to the type of variety, as L1 languages seem to function differently from L2 varieties. Substrate influence, on the other hand, cannot be the reason for the difference between the variety types, since all L2 languages behave very similarly despite their having substrates from various language families with grammatically extremely different structures. A more

121 The respective Chi square results for tense (df = 1) are as follows. British English: $\chi^2 = 13.274$, American English: $\chi^2 = 2.259$, Jamaican English: $\chi^2 = 0.103$, Indian English: $\chi^2 = 0.36$ and Singapore English: $\chi^2 = 0.435$. The critical level of 2.706 for $p = 0.1$ is exceeded by the result for British English only.

122 Note, however, that assumptions relating to American English have to be made very carefully, as the insignificant Chi square test seems to contradict the rather obvious difference in percentages between non-agreement rates in present tense as opposed to past tense sentences. For this reason the postulation of a potential difference in patterns correlating with the type of variety should be understood as a very tentative hypothesis whose validity needs to be tested in further investigation.
thorough investigation of a larger number of L1 and L2 varieties would be necessary to test whether this assumption is valid. The next grammatical feature to be discussed is contractedness.

### 5.4.5 Contractedness

The investigation by Martinez and Martinez (2003) and several other linguists interested in existential *there* + *be* constructions indicates that contractedness exerts a strong influence on subject-verb agreement in such existentials. The following discussion of the influence of contractedness on non-agreement in existential *there* + *be* constructions is based on the conversational category S1A only, as the tendencies found are more pronounced in the informal data than in the whole spoken corpus sections. This is most likely due to the higher overall number of contracted forms in informal private language than in more formal public speech, which greatly facilitates drawing reliable conclusions about contractedness as a potential factor influencing agreement in existential *there* + *be* constructions.

Only positive sentences, such as (15) are analysed, while negative sentences such as (16) were not counted in the analysis in order to avoid a bias resulting from differences relating to polarity rather than contractedness.

(16) I mean let's ha in if **there isn't such people** like Haringey there wouldn't be people like me (ICE-GB)

Plural contracted verb forms, such as the one presented in example (17), are not very frequent in my data, with only eight occurrences in the British data, three in the American sub-corpus, 26 in Jamaican English, nine in the Indian data and 15 in the corpus of Singapore English.

(17) I mean on a day like this it's magnificent and **there're some really nice people** there (ICE-GB)

Non-agreement involving such contracted plural verbs in sentences with singular subjects is hardly attested at all. It occurs only twice in Indian English (an example is given as 18) and three times in Singapore English (19).

(18) Sandhyanand **there're some clipping** from Sandhyanand (ICE-India, S1A-099)

(19) He said that if **there’re any good play** call him (ICE-Sing, S1A-079)
In both cases, the non-agreement might be a result of a lack of plural inflection in the noun rather than a number mismatch between the subject and its verb. However, as it is not possible to determine whether the nouns were intended as plural or not, the overt structure is what counts and the examples had to be regarded as instances of non-agreement. In order to be able to compare my results to those reported in previous investigations, my analysis has to be restricted to *there* + a singular verb form of *be*. Figure 23 gives the rates of non-agreement in existential *there* + *be* constructions in the category S1A for full forms such as *there is* and contracted forms such as *there’s* in sentences with plural subjects.

![Figure 23: Non-agreement with full versus contracted verb forms in category S1A](image)

Figure 23 reveals that there is more non-agreement when the verb form is contracted, as in (20), than with a full verb form (21).

(20) **There’s** always two guys in there (SBC)

(21) Well **there** is a lot of changes (ICE-JA)

While this tendency is visible in all varieties under consideration, the difference in non-agreement rates is much higher in the native varieties British and American English than in the postcolonial Englishes from India and Jamaica. The rate of non-agreement in sentences with plural subjects involving the full form *there is* appears to be highest in American English with 28.9%, but this is possibly due to the corpus compilation, and otherwise ranges between 3.4% in British English and 12.0% in Singapore English. The rates are thus
relatively low and there is no clear difference with regard to variety type. By contrast, when the verb is contracted the rate of singular agreement with plural subjects increases dramatically in the native varieties. Default singulars, i.e. the use of the contracted form *there’s* with plural subjects, can definitely be regarded as the preferred form in my data for British and American English, with 88.7% and 92.9% respectively, while the sub-corpora for the non-native varieties Jamaican English (13.3%) and Indian English (18.2%) do not show this strong preference for default singulars, as the non-agreement rates are only slightly higher in contracted than in full forms. In other words, the trend toward a preference for singular agreement is visible in these varieties, but it is rather weak.

In Singapore English, the rate of singular agreement, i.e. contracted *there’s*, with plural subjects lies somewhere in between those found in the native varieties and those in the postcolonial Englishes, but the patterns resemble the L1 varieties more closely than the other L2 varieties. This difference relating to variety type becomes even clearer when the differences in percentages are taken into account, as in Table 18.

<table>
<thead>
<tr>
<th></th>
<th>Non-agree</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full</td>
<td>Contracted</td>
</tr>
<tr>
<td>BritE</td>
<td>3.42%</td>
<td>88.73%</td>
</tr>
<tr>
<td>AmE</td>
<td>28.95%</td>
<td>92.86%</td>
</tr>
<tr>
<td>JamE</td>
<td>5.49%</td>
<td>13.33%</td>
</tr>
<tr>
<td>IndE</td>
<td>6.45%</td>
<td>18.11%</td>
</tr>
<tr>
<td>SingE</td>
<td>12.04%</td>
<td>59.46%</td>
</tr>
</tbody>
</table>

Table 18: Non-agreement with full versus contracted verb forms in category S1A

The differences in non-agreement rates are greater in the native varieties British and American English (85.3% and 63.9%, respectively) and smallest in the non-native varieties Jamaican English and Indian English (7.8% and 11.7%, respectively), with Singapore English ranging in between (47.4%). Chi square tests of a correlation between contractedness and the amount of non-agreement in existential *there + be* constructions in each variety confirm this postulated difference, as the test is significant for British English, American English and Singapore English, but not for Jamaican and Indian English.

The results in Figure 23 thus support Meechan and Foley (1994) and Crawford (2005), who found contractedness to be a factor that exerts a strong influence on subject-verb agreement in Canadian English and American English, i.e. L1 varieties. My data also provide strong evidence for the postulation by Kortmann and Szmrecsanyi (2004) that
default singulars are a widespread feature that is attested in varieties of English around the world, but that its frequency distribution differs with variety type, i.e. feature [55] “is categorical in all L1 varieties” (Kortmann and Szmrécsey 2004: 1199), but not in L2 varieties (or pidgins and creoles). Although many studies into linguistic variation have found this use of non-standard –s in existential *there + be* constructions with plural subjects to be “one of the most widespread, albeit non-standard, features of English varieties” (Rupp 2005: 255), these ‘default singulars’ (cf. Chambers (2004) and Tagliamonte (2006)) cannot be considered the “proto-typical exemplar of a ‘vernacular universal’” (Tagliamonte 2006), because the strong preference for singular agreement in existentials, particular in contracted verb forms, is restricted to native varieties.

Figure 23 indicates that the patterns in Singapore English differ considerably from those of the other non-native varieties. However, the fact that the pattern is very similar to those found in the native varieties practically rules out substrate influence as a potential explanation. For substrate influence to be plausible, the patterns in Singapore English would have to differ not only from the other non-native varieties but also from the native ones where no substrate is present that might exert an influence on the patterns. This, however, is not the case. Moreover, the substrate languages present in Singapore are isolating languages, similarly to the creole in Jamaica, and they differ greatly from the substrates in India. Nevertheless the patterns in Jamaican and Indian English are very similar indeed, so that influence from those very different background languages cannot be responsible for the frequency patterns found. The question of what qualifies as a possible explanation of the similarities between Singapore English and the native varieties as well as some hypotheses as to the nature of potential substrate influence will be addressed below in chapter 6, after the analysis of polarity as another linguistic constraint on non-agreement in existential *there + be* constructions.

### 5.4.6 Polarity

In previous research on agreement with the verb *be*, polarity has been found to influence agreement patterns in various native Englishes (cf. section 3.2 and the summary of previous work by Walker (1997)). Examples for non-agreement are presented as (22) for positive polarity, and as (23) for negative polarity.

(22) The rules actually say that if *there're* any **protest** the only people who… (ICE-Sing)
(23) **There’s no side effects** of homeopathy (ICE-Ind)

Figure 24 indicates that in all varieties investigated, non-agreement is more frequent in positive than negative sentences.\(^\text{123}\)

![Figure 24: Non-agreement by polarity](image)

In British English, 7.4% of all positive text units have non-standard agreement, while only 4.4% of negative text units are cases of non-agreement. The numbers for American English are 18.6% non-agreement in positive versus 9.8% in negative sentences, for Jamaican English 5.2% versus 3.3%, for Indian English 6.1% versus 4.0% and, last but not least, for Singapore English 7.0% versus 6.4%, respectively.

Although this tendency for more non-agreement in positive as opposed to negative sentences is only very slight in my data, this finding ties in with the observations made by Martinez and Martinez in a corpus of British English, who report that “TCs [i.e. *there* constructions, S.J.] showing non-concord present lower frequencies of negative polarity, especially in the case of speech” (2003: 270). As one can see in Figure 24, the difference in rates of non-agreement in positive versus negative sentences appears unremarkable in all ICE-corpora, with 3.1% in BritE, 1.9% in JamE, 2.1% in IndE and in 0.6% SingE. By contrast, the Santa Barbara Corpus reveals a much larger difference for American English, namely 8.8%\(^\text{124}\). However, Chi square tests for polarity and the amount of non-agreement

---

\(^{123}\) The analysis of the effects of polarity on non-agreement in existential *there + be* constructions is based on the whole spoken corpus sections. The percentages are to be understood as proportions of the total of positive versus negative text units.

\(^{124}\) Whether this large difference in American English in comparison with the other varieties is a consequence of the restricted data cannot be determined here, as a truly comparable data basis would be necessary.
reveal that this tendency is so slight as to be insignificant in all varieties except British English.\textsuperscript{125}

Previous research in native varieties of English suggested that non-agreement in sentences with negative polarity occurs particularly frequently when the determiner *no* is used (Rupp 2005: 262f., in her investigation of non-standard –s in expletive *there* sentences in Burntwood, GB, gives a summary of results found in previous studies, among them Meechan and Foley (1994) for Canadian English, Britain and Sudbury (2000) for Falkland Island English and New Zealand English).

Although the rates of non-agreement in sentences with negative polarity are very low, ranging from 3.3\% in Jamaican English to 9.8\% in American English, these cases can nevertheless indicate whether the determiner *no* indeed favours non-standard agreement in existential *there* constructions in my data. Table 19 gives the frequencies of non-agreement with particular types of negation. The category ‘No’ includes sentences like (24) in which the negator *no* is used, ‘Not’ comprises sentences with full verb forms + *not*, e.g. *is/are/was/were not*, such as the one presented in (25), the category ‘Contracted’ subsumes all sentences with the contracted forms *isn’t / aren’t / wasn’t or weren’t* (e.g. example 26), and ‘Others’ includes all remaining cases which cannot be classified as any of the above, for instance (27).

(24) No I mean *there’s no seats* left on that day (ICE-GB, S1A)

(25) …um if it’s a bad year *there is just not a lot of food resources*. (SBC)

(26) Since I have switched from linguistics to language communication and society *there isn’t much uhm field trips* within that area (ICE-Jamaica, S1A-057)

(27) He maintains his point though perhaps in a strong and hard way that *there was never any conditions* attached to the property being transferred to him (ICE-Singapore, S1A-068)

\textsuperscript{125} The Chi square results are as follows. British English: $\chi^2 = 3.382$ (p = 0.066), American English: $\chi^2 = 2.373$ (p = 0.12), Jamaican English: $\chi^2 = 1.328$ (p = 0.25), Indian English: $\chi^2 = 2.053$ (p = 0.15) and Singapore English: $\chi^2 = 0.132$ (p = 0.72). The fact that the 8.8\% difference in non-agreement rates in American English between sentences with positive as compared to negative polarity nevertheless yields an insignificant Chi square test is most likely due to the much smaller data base.
<table>
<thead>
<tr>
<th>Negation Type</th>
<th>BritE</th>
<th>AmE</th>
<th>JamE</th>
<th>IndE</th>
<th>SingE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tokens</td>
<td>%</td>
<td>Tokens</td>
<td>%</td>
<td>Tokens</td>
</tr>
<tr>
<td>No</td>
<td>11</td>
<td>91.7</td>
<td>1</td>
<td>20.0</td>
<td>4</td>
</tr>
<tr>
<td>Not</td>
<td>0</td>
<td>0.0</td>
<td>2</td>
<td>40.0</td>
<td>1</td>
</tr>
<tr>
<td>Contract.</td>
<td>1</td>
<td>8.3</td>
<td>2</td>
<td>40.0</td>
<td>1</td>
</tr>
<tr>
<td>Others</td>
<td>0</td>
<td>0.0</td>
<td>0</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>100</td>
<td>5</td>
<td>100</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 19: Non-agreement by negation type

Table 19 shows that, in four out of the five varieties covered here, non-agreement in negative sentences occurs most frequently with the determiner *no*. American English is the only variety in which non-agreement is less frequent with *no*-negation than with other negation types. However, with just five instances of non-agreement in sentences with negative polarity in the sub-corpus of American English, the token numbers are too small to be significant, i.e. a larger sample would be necessary to test the reliability of this result. In the other four varieties, more than half of the cases of non-standard agreement in negative sentences involve the determiner *no*, the percentages lie between 57.1% in Jamaican English and 91.7% in British and Indian English. The postulated influence of the type of determiner in the post-verbal noun phrase on non-agreement in *there + be* constructions (see Rupp 2005: 262) is thus confirmed by my data not only for native varieties but also for the L2 Englishes considered here.

5.5 Summary

As the results show, non-agreement in existential *there + be* constructions is not very frequent in any of the varieties, although it turned out to be slightly more frequent in native than non-native Englishes. The fact that non-agreement in the most informal category of private conversations is highest in British and American English may be a consequence of the fact that in ESL contexts, for instance in India and Jamaica, English tends to be used for rather formal linguistic purposes such as administration and commerce, while for private informal communication other codes are the preferred choice, e.g. lower mesolectal or basilectal varieties of the creole continuum (i.e. Patwa) in Jamaica and one of the many local languages in India. Thus the use of English automatically means a more formal language usage in these varieties. The frequencies of non-agreement found in the sub-corpus of Singapore English lie between those of the native varieties and the postcolonial...
Englishes, but in the analysis of all grammatical factors investigated they resemble those of the L1s more closely than those of the L2s.

In the analysis of the four ICE-corpora and the Santa Barbara Corpus of spoken American English the following tendencies were found in all varieties:

1) Variable agreement—in none of the varieties under investigation is agreement or non-agreement categorical. Instead, the majority of verb forms are chosen in accordance with the grammatical number of the post-verbal noun phrase, but non-standard agreement does occur, albeit not frequently, in all varieties.

2) Register variation—there is considerably more non-agreement in the most informal text type than in the most formal category. The amount of variation is greatest in the most informal category of conversations, where the native varieties and Singapore English have the highest rates of non-agreement.

3) ‘Default singulars’—in all varieties,¹²⁶ there is more non-agreement when the post-verbal noun phrase is plural than when it is singular. Subject number is thus a strong predictor of non-standard agreement, especially in conjunction with contractedness. This tendency is stronger in the native than in the non-native varieties.

4) Contractedness—in the L1 varieties and Singapore English, the contracted form displays non-standard –s significantly more frequently than the full form of the verb. This contrasts with Jamaican English and Indian English, where there is no statistically significant difference relating to contractedness of the verb form.

5) Tense—the statistical test of tense reveals that any apparent differences are insignificant in most of the varieties under consideration. Only in British English is there a significant influence of tense in that non-agreement is more frequent in sentences with present tense rather than past tense reference.

6) Polarity—the tendency of non-agreement to occur more frequently in sentences with positive polarity is visible in the data for all native and non-native varieties considered. However, Chi square tests indicate that this tendency is so slight as to be insignificant in all varieties except British English. If the instances of non-agreement in negative sentences are analysed separately, the results show that substantially more non-agreement occurs when the determiner no is used than when the sentence is negated by other means. This is true for all Englishes except American English, but the numbers for that variety are too small to allow generalizations.

¹²⁶ Note, however, that the Chi square test shows that this difference relating to subject number, though visible in the data, is not significant in Indian English.
After this analysis of the grammatical factors influencing agreement patterns in existential *there + be* constructions in the five varieties of English, the next section will be devoted to a description of the typology of existential constructions, before the particular ways of expressing existence employed by speakers of the substrate languages are explored. Then a corpus analysis of the informal spoken sections of ICE-corpora, i.e. the text category S1A comprising casual conversations and phone calls, will provide a clearer picture on the question to what extent and how substrate influence and universal tendencies contribute to the complex patterns of existential constructions in post-colonial contact varieties of English, particularly in Jamaican English.
6 Existential Constructions

6.1 Expressing Existence

A typical example of an existential construction in English as discussed at length in the previous chapter is given as (1a).

(1a) There is a dog in the garden.\textsuperscript{127}

In existential sentences, some indefinite entity or ‘theme argument,’ here a dog, is related to some location or ‘location argument,’ here the garden by means of a copular verb, usually a form of be (cf. Freeze: 2001: 941). In English, indefiniteness of an entity is expressed by an indefinite determiner such as a/an or the absence of a determiner in case of plural nouns, as in (1b).

(1b) There are dogs in the garden.

A peculiar but typologically anomalous (cf. Freeze 2001: 941) characteristic of English existentials is the dummy subject there occupying the syntactic position of the subject. Freeze argues that there is substantial variety across the world’s languages with regard to constructions expressing existence (2001: 941) but demonstrates that such “proform existentials” (p. 944) exist only in a few Romance and Germanic languages, Arabic and some Austronesian languages and that these constitute only “a distinct minority of languages” (p. 944). Freeze reports that the locative proform is not attested in languages with the typical word order SOV, e.g. Tamil or Hindi, and that in the languages where it does occur it “obeys strict positional constraints” (2001: 944). Among others he gives the following examples from French (2) and Italian (3) (Freeze 2001: 941).

(2) \textit{il y a des chocolats sur la table}

\begin{tabular}{c}
3sg there have some chocolates on the table
\end{tabular}

‘There are some chocolates on the table.’

(3) \textit{ci sono uomini nella casa}

\begin{tabular}{c}
there are men in the house
\end{tabular}

‘There are (some) men in the house.’

\textsuperscript{127} Sentences (1a and b) are constructed examples used for illustration. Examples (2) through to (4) are taken from Freeze (2001).
The proform in these examples always precedes the locative phrase. Freeze states that such proforms are generally locative, except in some Germanic languages, which instead have non-locative expletive pronouns in subject position (2001: 949, but see also Freeze 1992), for instance Swedish *det* or German *es* (4).

(4) *es* gibt/ist *ein* Buch *auf* dem Tisch

it gives/is a book on the table

‘There is a book on the table.’

Typological research, for instance by Stassen (2008b) and Lyons (1967, 1968), has shown existentials such as (1a, repeated below) to be closely related to locative sentences (5) (cf. also Dryer 2007: 241 and Schachter and Shopen 2007: 56).

(1a’) There is a dog in the garden.

(5) The dog is in the garden.

Clark (1978) notes that existentials, locatives and two types of possessive sentences, the *have*-possessive (6) and the *be*-possessive (7), have the same surface constituents.

(6) John has a dog.

(7) The dog is John’s.

Following the hypothesis put forward by Lyons (1967) that “both the existential and the possessive construction in each language are derived from the same source, namely from the locative” (Clark 1978: 90), Freeze suggests that all four of these constructions can be structurally related as “variations on a single underlying array of elements” and thus can all be accounted for by what he calls the ‘locative paradigm’ (2001: 946).

As Clark emphasizes, the locative source of the existential construction, while not always visible in the surface structure, is inherently present in all existential constructions. In support of this claim he reports findings by Hintikka (1968 and 1968-69) expressing the same notion, namely that “for an object to be (i.e. to exist) normally means that it is to be found somewhere in space” (Clark 1978: 89). For this reason, Clark adds, the existence of some entity should be expressed in locative terms (1978: 89).

In the existential (1) and the locative construction (5) above there is obviously a locative element, *in the garden*, as well as a nominal which functions as the theme. What mainly distinguishes these two sentences is the word order in which these constituents are
positioned in the sentence, and the fact that the nominal is indefinite in (1), a dog, but definite in (5), the dog. Generally, definite articles are used to indicate that the item referred to has already been mentioned before, i.e. it is known information, whereas indefinite determiners signal that the item is being introduced as new information. Clark (1978: 88) therefore concludes that “the absence of indefinite nominals in initial position reflects a general discourse constraint in languages” which states that “[n]ew information is typically introduced after given information.” The non-deictic there in English existential constructions is thus claimed to have the function of filling the initial position in order to “allow […] the introduction of new material in second position” (Clark 1978: 88).

Dryer (2007) considers definiteness to be a central element in the relation of locative and existential expressions and illustrates this with three examples (8) – (10) from Ma’anyan, an Austronesian language spoken in Kalimantan (Borneo).

(8) inehni naqan hang sungking
    mother be.at at kitchen
    ‘his mother is in the kitchen’

(9) naqan erang kaulun wawey mawiney hang tumpuk y eruq
    be.at/exist one CLSFR woman beautiful at village the
    ‘there was a beautiful woman in the village’

(10) sadiq naqan tumpuk eteqen
    olden.time exist village Eteen
    ‘once upon a time there was a village called Eteen’ (Dryer 2007: 240f.)

Although all three of these expressions use the locative copula naqan ‘be at, exist,’ they are neither all locative constructions nor are they all existential expressions. While in (8) and (9), the copula naqan links a theme argument (his mother versus a beautiful woman) to an expression denoting a location (in the kitchen versus in the village) the same is not the case in (10), which has no overt locative element. Thus, according to Dryer (2007: 241), (8) and (9) can be considered locative expressions but (10) cannot. Similarly, only in (9) and (10) does the copula state the existence of something (a beautiful woman versus a village called Eteen), but this aspect is absent in (8), as the existence of his mother is presupposed. That is, (9) and (10) are existential constructions, but (8) is not. Dryer

---

128 Dryer regards only those constructions as locatives which have overt locative complements. He does not take into consideration potential underlying locative notions of ‘existing in space’ (cf. Clark 1978: 89) which are allegedly present in all existential constructions (cf. above).
concludes that the contrast between (8) and (9) “thus corresponds to a pragmatic difference of identifiability, and hence to a grammatical difference in definiteness in English” (2007: 241). He emphasizes that “[m]any languages are like Ma’anyan and English in using two different constructions with locative predicates depending on whether the theme is identifiable [i.e. definite, S. J.] or not, with a distinct existential construction being used when the theme is nonidentifiable” (Dryer 2007: 242), i.e. indefinite.

An important point to note is that languages around the world differ greatly with respect to how they signal definiteness and indefiniteness, i.e. ±Definite. Clark (1978: 91) argues that in contrast to English, which has both indefinite and definite determiners, many languages only use one definiteness marker, usually +Definite (Moravcsik 1969), some have an optional marker for indefiniteness, and others do not have definiteness markers at all. In languages of the last group, the main indicator of definiteness is word order (Clark 1978: 91).

In contrast to existential and locative constructions, the locative element is not so obvious in case of the possessives presented as (6) and (7). Clearly both sentences contain a noun referring to the item possessed, i.e. the indefinite nominal a dog in (6) versus the definite noun phrase the dog in (7). Only definite nominals usually occur in sentence-initial position for the same reasons as above.

As concerns the locative origin of such possessive constructions, Clark argues that “the possessor in the two possessive constructions is simply an animate place” (1978: 89), i.e. that “[t]he object possessed is located in space. […] but that] the place happens to be an animate being” (p. 89). The same is demonstrated by Freeze with examples from numerous languages, which show that “not only is the possessive sentence an existential, but in all [examples], the semantic relationship of the location and the theme arguments is the same” (Freeze 2001: 943, cf. also Freeze 1992 for a further elaboration on this point). He consequently comes to the conclusion that “existentials and possessives are the same, except for the semantic features of the subject” (2001: 944), i.e. the [± human] feature.

Another piece of evidence for the close relation between existential constructions and possessives comes from the fact that many languages around the world use the same copular verbs for expressing possession and existence, e.g. Hindi hōnā and Mandarin Chinese you. Her (1991: 383) attributes this to the fact that relations of possession are at

---

the same time relations of existence, as “A possesses B is in effect equivalent to that B exists within the domain of A” (Her 1991: 383).

An important point to note here is that possession is not necessarily to be understood as ownership. According to Her (1991), ownership is a specific kind of possession, which in turn is a specific kind of existence.

Let us look at sentences (6’) and (11) to (13).

(6’) John has a dog.
(11) I have three sisters.
(12) Brazil has beautiful beaches.
(13) The year has twelve months.

Only the first of these examples can be interpreted in terms of ownership, i.e. the possessor John is in fact the owner of the possessed entity a dog. By contrast, the other three examples are not instances of ownership but are better interpreted as expressing existence. That is, what (12) expresses is not that Brazil owns beautiful beaches, but that the possessed items, i.e. beautiful beaches, exist within the domain of the possessor, i.e. Brazil.

In English and some other languages that use different verbs to express existence and possession, this close relation between these constructions becomes even more evident when one takes into account that existential constructions can be transformed into possessive structures and vice versa. To give an example, when the existential sentence (14)¹ is transformed into (15), the locative item, here the table, is made the subject of the sentence and verb changes from be to have.

(14) There is a book on the table.
(15) The table has a book on it.

As the locative now precedes the theme, a copy on it is obligatorily added which follows the theme (cf. Clark 1978: 109). The result of this transformation is then a possessive of the type exemplified in (11) to (13), to be interpreted in terms of existence rather than ownership.

According to Freeze (2001: 950), there are only a few languages in the world that use have-possessives along with be-possessives and many more that use only one copular verb for both types of possessives. The few languages that do use have include some Indo-

¹ The examples from Clark (1978: 109) are used here for illustration.
European languages, for instance Germanic (German, English), Romance (Portuguese, Spanish) and Persian (cf. Freeze 2001). The following sentence is an example from Portuguese taken from Freeze (2001: 951), in which the same proposition is expressed by means of a *be*-possessive in (16a) and a *have*-possessive in (16b).

(16a) *O menino está com fome*

The child is with hunger

‘The child is hungry.’

(16b) *O menino tem fome*

The child has hunger

‘The child is hungry.’

Both versions are commonly used possessive constructions in Portuguese. They both contain an animate (+ human) noun in subject position, *o menino*, and an inanimate and indefinite noun *fome*, which functions as the theme. What is different is that in (16a) the possessor is presented as the locative element whereas the same is not true in (16b). Not only can locative constructions be used as possessives, but possessives can also be used as existentials, as the following Portuguese example with the verbs *haver* ‘have’ and *ter* ‘hold, grasp’ shows.

(17a) *Há muitas praias no Brasil.*

It has many beaches in Brasil.

‘There are many beaches in Brazil.’

(17b) *Tem muitas praias no Brasil.*

It holds many beaches in Brasil.

‘There are many beaches in Brazil.’

---

131 Stassen (2008a) notes that 63 out of 240 languages use *have*-possessives to express the possession of an entity. However, his focus does not lie on possessive constructions that can function as existential expressions, i.e. all four subtypes discussed above (6) and (11) to (13) and the co-occurrence of them with *be*-possessives, but on those expressions “which encode the concept of alienable possession”, i.e. ownership. That is, he includes only the first subtype of possessives, as in (6). Stassen’s map in the *World Atlas of Language Structures Online* (feature 117) indicates that the distribution of constructions used to express alienable possession “shows considerably areal effects” (Stassen 2008a). For example, most of the languages in India, including Hindi and Tamil, use locational constructions, while languages in East and Southeast Asia, e.g. Mandarin Chinese, prefer topic possessives, and central and western European languages such as English use mainly *have*-possessives.
Another example comes from German, where *es gibt* (18a) and *es ist/sind* (18b) are the most common forms of the existential, but in some dialects a form with *haben* ‘have’ (18c) can be used (see Elspaß and Möller 2008 for more information on existential constructions in German). The English translation in each case is ‘There are many hotels here.’

(18a) *Es gibt hier viele Hotels.*
   It gives here many hotels.

(18b) *Hier sind viele Hotels.*
   Here are many hotels.

(18c) *Hier hat es viele Hotels.*
   Here has it many hotels.

The last type of existential to be mentioned here is what Freeze calls the *possessed-theme possessive* (2001: 951). He lists this as another locative structure in many of the world’s SOV languages, e.g. Mayan, Turkish, Hindi and Hungarian. Freeze reports that languages differ with regard to the existential forms that occur, in that “SOV languages (e.g. Japanese) have no proform existentials, and SVO languages (e.g. English) have no possessed-theme existentials” (2001: 952). He presents an example for a possessed-theme possessive from Hindi (cf. Freeze 1992), which is reproduced as (19) below.

(19) *meree doo bhaii hai*
   my.PL two brother COP.3 PL
   ‘I have two brothers.’ (Literally: My two brothers are.) (Freeze 2001: 952)

Freeze shows that this construction has a locative copular verb, the possessor is marked for genitive case according to the characteristics of the particular language, and the possessed item functions as the subject (2001: 951). The next section briefly explores which of the constructions discussed here can actually be found in the background languages of Jamaica, Singapore and India.

---

132 The examples were taken from the third round of a research project on regional uses of lexical and grammatical items in German supervised by Elspaß and Möller since 2003. For more information consult the website of the *Atlas zur Deutschen Alltagssprache (ADA)* at <http://www.philhist.uni-augsburg.de/de/lehrstuehle/germanistik/sprachwissenschaft/ada/>
6.2 Existential Constructions in the Substrates

In order to find out how contact between Jamaican English and its creole substrate might be reflected in existential constructions found in the corpus, it is necessary to look at how the existence of some entity is typically expressed in Jamaican Creole. According to Patrick, “JC [Jamaican Creole, S. J.] normally expresses existence with the invariant verb (h)av plus an indefinite pronoun subject, e.g. dem ‘they’ or yu ‘you’” (2007: 140). In other words, existence is most commonly expressed by means of possessive constructions involving the possessive copular verb corresponding to English have, i.e. the first of the two types of possessive constructions. Examples of such structures are given as (20), where dee hav represents the indefinite pronoun subject plus invariant verb, and (21), where y’av fulfills that function.

(20) *dee hav a gruup a man niem Stepaz*
   they have a group of men called Steppers
   ‘There was a group of men called the Steppers’ (Patrick 2007: 140)

(21) *y’av a glas chrch hAp die*
   you have a glass church up there
   ‘(There is) a glass church up there’ (Patrick 1989, line 25)

In his summary of grammatical features of Jamaican Creole Patrick reports that, “occasionally, existence is also expressed with locative copula de” (2007: 140) and gives (22) as an example.

(22) *Mi no biliib nobadi de*
   I not believe nobody there
   ‘I don’t believe there’s anybody’ (Patrick 2007: 146)

Proceeding now to the languages in Singapore, Leimgruber (2008) notes that in Singapore English a possessive form with got is commonly used to express existential functions, as shown in example (23).

(23) *Got books on the table.*
   ‘There are books on the table.’ (Leimgruber 2008)

This construction is commonly used without an expletive or dummy subject in Singapore English, probably due to a tendency in Mandarin Chinese to drop subjects under certain
circumstances (cf. Huang 1989), but previous published work does not provide enough information to generalize or explain which factors are responsible for this use of got, so that further research is necessary to validate this assumption. However, this got-existential has been identified as a distinct marker of the colloquial variety of Singapore English (see Gupta 1994: 10f. and Teo 1995), occurring mainly in casual conversations, that is, in registers in which influence from substrate languages is possible and most likely to occur. The main background languages in Singapore as described above are Mandarin Chinese and other Chinese dialects, Malay and Tamil. That it is common in Mandarin Chinese to realise existential statements by means of possessive constructions becomes clear from the observation that “Chinese you 'have', when construed as existential, is traditionally analyzed as a modal verb or an auxiliary“ (Tsai 2003: 1). Similarly, Clark notes that Mandarin Chinese uses the same verb form, you, for existentials and possessives (1978: 104f.), but a different verb for locatives, namely shì, or no verb at all (1978: 103). According to Omar, “[i]n the Western Austronesian languages the verb for ‘have’ is the same verb which denotes existentiality” (1974: 391). In other words, existence is expressed by means of a possessive construction. In addition to this multifunctional verb, there is another verb in Malay that is restricted to the possessive meaning of ‘have’ (cf. Omar 1974). Tamil, which is spoken in Singapore as well as in India, does not apply possessive constructions for existential functions but uses one of two verbs, the invariant verb uṇṭu ‘exist’ (24), or irukka ‘be located’ (25).

(24) kaṭavu/ uṇṭu
god-NOM exist
‘God exists,’ ‘There is a God.’ (Steever 2005: 170)

(25) kaṭavu/ ??(arkē) irukkipāṇ

god-NOM there be-PRES-2SG (Steever 2005: 171)

Examples of the first kind typically occur when the existence of an entity is explicitly expressed, while the second is frequently used for sentences with overt locative expressions. This type of expression “thus appears to coordinate spatial and temporal dimensions: one is in a place at a time” (Steever 2005: 170). While Lindholm (1969: 6) argues that irukka is best used to express the existence of an entity, Steever disagrees and
proposes that *ūṇtu* “more properly serves this function” (2005: 170). In any case, Tamil does not commonly apply possessive constructions to express existence. Apart from Tamil, the background languages most widely spoken in India are Hindi, Bengali, Telugu and Marathi. However, as information on existential constructions in most of these languages is rather difficult to obtain, the only language that can be mentioned here is Hindi.

In contrast to English, proform locatives are not applied in Hindi to express existential functions. In their work on divergence in machine translation from English to Hindi and vice versa, Sinha and Thakur (2005) describe the divergence between these two languages in the mapping of existentials. They report that

Hindi does not have a pleonastic subject construction and the contrast between existential and non-existential (mostly definite) sentences is realized by several other ways [rather than there- and it-sentences, S.J.] such as the movement of the noun phrase from its canonical position and the use of demonstrative elements. (Sinha and Thakur 2005: 249f.)

The authors use examples (26a) and (b) to illustrate this use of word order, as these sentences differ exclusively in the respective positions of the subject NP and the local adverbial phrase.

(26a) *jangal meN sher hE.*

Forest in lion be.PR

‘There is a lion in the forest.’

(26b) *sher jangal meN hE.*

Lion forest in be.PR

‘The lion is in the forest.’

While (26a) is a typical existential construction with the word order locative – theme nominal – verb, the word order is different in the locative sentence (26b), namely theme nominal – locative – verb (cf. Clark 1978 96). That Hindi makes extensive use of word order becomes even more evident where possessives are concerned. As Freeze (2001: 952) demonstrates with the following examples, the language has both the possessed-theme existential (27) and locative-subject existential (28).

(27) *meree doo bhaii hāi.*

my.PL two brothers COP.3 PL

‘I have two brothers.’ (Literally: My two brothers are.)
(28) *la*kee-kee paas *kutta* hai

boy.OBL-GEN by dog COP.3 SG

‘The boy has a dog.’

The difference in word order of the constituents is related to whether the possessive is to be interpreted as ownership or “alienable possession” (Freeze 2001: 952), e.g. (28), or in terms of existence, i.e. “inalienable possession” (27). In other words, “in Hindi, inalienable possession has a theme subject, while alienable possession has a locative subject” (Freeze 2001: 952).

Following Freeze’s account that SOV languages do not have proform existentials and that possessed-theme existential do not occur in SVO languages, we can note that of the many background languages in India, all of which have the typical word order SOV, none have proform existentials, while the possessed-theme existential is largely absent from the background languages in Singapore, most of which are SVO languages.\(^{133}\)

In sum, there are several types of constructions that are of interest in the analysis of substrate influence in the varieties of English spoken in Jamaica, Singapore and India. The first type comprises locative constructions with *there* in post-copular position (attested for instance in Jamaican Creole) or in pre-copular position (i.e. the typical English *there + be* pattern analysed in detail above, in which *there* is used non-deictically as a dummy subject).\(^{134}\) Other expletive subjects will also be analysed when co-occurring with *have*, for instance *you* or *them*, as these have been identified as commonly expressing existence in the Jamaican substrate. Moreover, the use of word order will be considered due to it being a crucial feature distinguishing between existentials and locatives as well as between different types of possession, for instance in Hindi. The last type of construction frequently used to express existence is possessives. This group is sub-divided into *have*-possessives, *got*-possessives and *be*-possessives, the latter of which will, however, not be analysed in the corpora, as these cases are difficult to extract by means of the WordSmith concordance program.

---

\(^{133}\) The one exception is Tamil, which has the word order SOV and is spoken in both, Singapore and India.

\(^{134}\) Constructions with expletive *there* in pre-verbal position have been analysed in much detail in chapter 5 and will therefore not be discussed again here.
6.3 Existential Constructions in Varieties of English

In order to analyse the different types of existential constructions and possessive constructions in the native variety British English\(^{135}\) as well as the non-native varieties Jamaican English, Indian English and Singapore English, a corpus analysis was carried out in each of the corpora by means of the WordSmith concordance program. The collocations of the following two lexical entries were searched: THEY*, yielding all instances of *they got, they have, they have got, they’ve and they’ve got* co-occurring with an indefinite noun phrase, **YOU*\(^*, **yielding all tokens of *you got, you have, you have got, you’ve and you’ve got* with indefinite noun phrases, but also instances of *them have, it have\(^{136}\) and zero subject + got. Moreover, the use of locative constructions for existential functions was analysed by means of a corpus search of *is there, are there, was there and were there*, in order to find cases where locative adverbial *there* was used in place of dummy *there*. Excluded from the analysis were questions, tag questions and unclear cases. Agreement or non-agreement was not analysed, as it is the occurrence and frequency pattern of the structures rather than the grammatical properties that are presently of interest.

The analysis yields the following results. In the conversational category S1A of ICE-Great Britain, 56 tokens of possessive constructions with the subject pronoun they have been found, and the numbers in the corresponding sub-corpora of ICE-India are 66 tokens, in ICE-Jamaica 74 tokens and in the data from Singapore 112 tokens. In terms of total numbers, Singapore English is the variety with the highest number of possessives of this type.

The frequency distribution is clearest when only those instances are considered in which the indefinite noun phrase is singular and introduced by the indefinite article *a/an*. Figure 25 therefore shows only those collocations of possessive constructions in which indefinite articles are used,\(^{137}\) in terms of token numbers (left) and with regard to percentages of the total (right). As the collocations *they’ve a* and *they have got a* are largely absent from the data they do not appear in the figure.

---

\(^{135}\) American English will not be analysed in this section as the data basis is not comparable in size to the other four corpora.

\(^{136}\) The form *them have* is relevant for this investigation as it is assumed to occasionally replace the Jamaican Creole construction *dem (h)av*. The expression *it have* was also included because it is rather common in Trinidad (Mair 2009, p.c.) and it will be interesting to see whether it also occurs in Jamaican English.

\(^{137}\) The total numbers of such instances counted are as follows. BritE: 15, IndE: 14, JamE: 24 and SingE: 30.
A search for the possessive construction *they have a* (or *an*) + noun phrase in the conversational category S1A of ICE-Jamaica yields 23 instances, an example of which is given as (29).

(29) They have a Burger King there. (ICE-JA, S1A-083)

By contrast, a search for the same string in ICE-GB yields only 9 tokens in the comparable portion of the corpus, while 25 tokens are recorded in the category S1A of ICE-Singapore and 12 cases are found in the same text type in ICE-India. The collocation *they’ve got a* + NP, as in (30), is most frequent in British English, with 6 tokens, and hardly attested at all in the non-native varieties.

(30) They’ve got a thing which is the equivalent to our Aga (ICE-GB)

There are very few instances of *they got a* + NP, e.g. (31), in my data, most of which were found in the Singapore sub-corpus.

(31) They got a double loop (ICE-Sing, S1A-085)

Although not all of the instances of *they have a* + NP are actually used for existential functions in the data, the numbers nevertheless indicate that there is a significant difference between the varieties with regard to the frequencies of such constructions, particularly between British and Indian English on the one hand (9 and 12 instances in S1A, respectively) and Jamaican and Singapore English on the other hand (23 and 25 cases, respectively). This difference does not seem to be related to the type of variety, since not all post-colonial varieties pattern similarly, and vernacular universals or angloversals thus
do not qualify as potential explanations. The more plausible explanation for this phenomenon is substrate influence. The high number of *they have a + NP* and related collocations in Jamaican English is here assumed to be the result of the tendency in Jamaican Creole to express existential functions by means of possessive structures (see Patrick 2007: 140). The equally high number of hits in Singapore English can be attributed to a similar tendency being at work in Chinese, one of the main substrates in Singapore (see the discussion above), a tendency, however, which is not present, or at least not as strong, in Indian English.

The possessive structures with the subject pronoun *you* will be analysed next, in order to find out whether this prevalence of possessive structures in Singapore English and Jamaican English rather than Indian English can be validated again in the sub-corpora of the four varieties. In the British sub-corpus, 86 possessive sentences with *you* as the subject where found, and the token numbers of the other corpora are as follows: 78 instances in the Indian English data, 162 in the Jamaican sub-corpus and 75 in the respective category in ICE-Singapore. In terms of total numbers, Jamaica stands out here in that it has considerably more such constructions than any of the other varieties. Figure 26 gives the token numbers of possessive constructions with the subject pronoun *you* (left) and the percentages of the total (right). The collocations *you’ve a* and *you have got a* are extremely infrequent in the corpora and therefore do not appear in the figure.

![Figure 26: Possessive constructions with the subject pronoun you](image)

---

138 Again, only those cases are counted where the noun phrase contains an indefinite article (*a/an*). The numbers of occurrences of such cases in the relevant sub-corpora are as follows. BritE: 32, IndE: 15, JamE: 52 and SingE: 21.
Examples of possessive structures with the collocations *you have a*, *you’ve got a* and *you got a* are presented as (32) to (34).

(32) … you **have a** visiting hours like from eight till one in the night (ICE-JA, S1A-080)

(33) You’ve just **got a** very slight uh enamel chipping there (ICE-GB, S1A)

(34) … and then usually winters now specially starts from November **you got a** different picnic seasons (ICE-India, S1A-008)

The figure reveals that, in contrast to all other varieties investigated, Jamaican English has a strong preference for expressing existence by means of the possessive collocation *you have a* + noun phrase. There are no instances at all of *you’ve got a* + noun phrase in the Jamaican and Indian data and very few in Singapore English. By contrast, British English employs a comparatively large number of these constructions. The patterns are reverse for the collocation *you got a* + noun phrase, which occurs most frequently in the Singapore data and is totally absent from the British sub-corpus.

From the two figures of possessive structures the following tendencies emerge: 1) Indian English is the variety in which the fewest possessive constructions were found, namely 15 involving the subject pronoun *they* and 14 involving *you* as the subject. 2) In terms of total numbers, British English uses relatively few possessive structures with the subject pronouns *they* and *you* but reveals the largest numbers of *they’ve got a* + noun phrase and *you’ve got a* + noun phrase. 3) Such Briticisms are totally absent from Jamaican English and rather infrequent in Singapore English and Indian English, in the latter of which these constructions occur only with the subject *they*. 4) The collocations *they got a* + noun phrase and *you got a* + noun phrase are totally absent from the British data and occur most frequently, but not in high numbers, in Singapore English.

Singapore English has the highest numbers of possessive constructions with the subject pronoun *they*, while Jamaican English reveals a strong preference for possessives of the type *you have a* + noun phrase. My data thus indicate that the postulated tendency in Jamaican Creole to express existence by means of possessive structures with the indefinite pronoun *you* (Patrick 2007: 140) strongly affects existential structures in Jamaican English by facilitating the use and thus causing the high numbers of *you have*-possessives in this variety (Figure 26).

In addition to the possessive structures described above, the Jamaican sub-corpus also contains examples of possessives with other subject pronouns, for instance one possessive
with *it* (35), and four tokens of possessive constructions involving the subject pronoun *them*, as in (36), the latter of which corresponds to the creole indefinite pronoun subject *dem* ‘they’ referred to by Patrick (2007: 140).

(35) Uh But *it has* so many opportunities (ICE-JA, S1A-028)

(36) I think *them* should realise *them have* a typhoid epidemic in Westmoreland (ICE-JA, S1A-097)

While constructions with the pronoun *it* also appear in the other post-colonial Englishes, albeit rarely, but not once in British English, possessives with the subject pronoun *them* are completely absent from all corpora except the Jamaican one. For this reason, the appearance of possessive structures with the pronoun *them* is a clear case of direct creole influence, attributable to the indefinite pronoun subject *dem* in Jamaican Creole (cf. Patrick 2007). As a direct creole construction, *them have* + indefinite noun phrase is highly salient and stigmatized and therefore does not occur frequently in the data. However, the expression *them have* functions as a bridge construction linking the creole form *dem have* with the standard English construction *they have*, and facilitating the use of the latter. This bridge pattern explains the high number of *they have* + indefinite noun phrase in Jamaican English as compared to British and Indian English.

As mentioned above, the corpus investigation revealed a prevalence of the collocations *they got a* + noun phrase and *you got a* + noun phrase in Singapore English. However, even more frequent in the Singapore data are sentences of the type illustrated in (23), reproduced below, in which *got* is used without a subject pronoun.

(23’) *Got books on the table.*

‘There are books on the table.’ (Leimgruber 2008)

In the conversational category S1A of ICE-Singapore, 46 instances were found, as compared to only one token in the corresponding portion of the British sub-corpus, two tokens in the Indian data and no attestation at all in the corresponding text category of ICE-Jamaica. Sentence (37) is the only construction of this type that was found in the British sub-corpus, and (38) and (39) are but two examples from Singapore.

(37) *Got* really good Christmas trees this year (ICE-GB, S1A)

---

139 Although it seems in this example that *it* refers to some previously mentioned noun phrase, the context of this conversation suggests that this is not the case. The construction *it has* or *it have* + indefinite noun phrase is, however, so rare in all varieties as to be negligible.
(38) So the people who were trained are youngsters *Got* some ang mos *Got* a lot of kwai lo uh (ICE-Singapore, S1A-025)

(39) Then *got* some very yuppie yuppie types (ICE-Singapore, S1A-025)

As discussed earlier, Indian English does not normally use possessive constructions in order to express the existence of some entity. There is, however, one type of construction that is much more frequent in the Indian data than in any other variety under investigation. The construction referred to is a seemingly locative construction used for existential functions, in which the common English dummy subject *there* is replaced by an adverbial *there*. Examples from ICE-India of such existential constructions with adverbial *is there, are there, was there* and *were there* instead of the English proform *there is/are/was/were* are presented as (40) – (43).

(40) **Competition is there** but this competition is not healthy often it is manipulated competition (S1A-089)

(41) I know in Karnataka I think all over Karnataka only **two schools are there** (S1A-063).

(42) Very horrible **dissection was there** (S1A-053)

(43) Some thirty thousand **species were there** (S1A-036)

Although none of these sentences has the canonical form with *there* in syntactic subject position, a subsequent form of *be* and—post-verbally—the notional subject, the examples are nevertheless used to point to the existence of the specific entity *competition, only two schools, very horrible dissection, and some thirty thousand species*, respectively. Only in Indian English is sentence-final *there* commonly used to express the existence of some entity. While this type of construction is very frequent in Indian English, with a total of 121 instances in category S1A of ICE-India, the numbers are extremely low in the other corpora. In British English, this construction was not attested at all, and its frequency in Jamaican English (one token) and Singapore English (two tokens) is negligible. In the latter three corpora, *there* generally functions as a locative adverbial when occurring in positions other than sentence-initially.

To sum up, Figures 27a-d give the four most frequently attested constructions\(^\text{140}\) in each variety that are normally used to express existence. The categories are to be understood as

\(^{140}\) The figure includes all types of (possessive and locative) constructions used for existential functions except the common English existential *there + be* construction discussed in detail in chapter 5.
follows: *they + have + n* includes all cases where the subject pronoun *they* co-occurs with a form of *have*, either in the full form *they have* or the contracted form *they’ve*, and an indefinite noun phrase that may, but need not, be introduced by the indefinite article *a/an*. The other categories are to be similarly interpreted.

**Figure 27a: Constructions used to express existence in British English**

**Figure 27b: Constructions used to express existence in Indian English**
The figures show that each variety of English has a clear preference for one or two particular constructions. While British English (Figure 27a) uses mostly possessives with the subject pronoun you and particularly you’ve got + indefinite noun phrase, this construction plays only a very marginal role in the other varieties, since it does not appear among the four most frequent constructions in any of figures representing the other varieties of English. By contrast, Indian English (Figure 27b) does not commonly rely on possessive structures at all for the expression of existence. In fact, Indian English is the only variety in which adverbial there is regularly used in place of dummy there, and the tendency to use this rather than possessive constructions is very strong. Although Sinha and Thakur state that “the contrast between existential and non-existential (mostly definite)
sentences is realized by several other ways such as the movement of the noun phrase from its canonical position and the use of demonstrative elements” (Sinha and Thakur 2005: 249f., cf. also section 6.2 above), this merely suggests that Hindi word order rules might exert an influence on existential constructions in Indian English, but it does not explain precisely how such rules in Hindi could account for the word order pattern found in the data of Indian English. In fact, on the basis of the research presented by Clark (1978: 96) one would expect existential sentences in Indian English to have the locative element in sentence-initial rather than in sentence-final position. It is unfortunately beyond the scope of the present work to explore in detail how this apparent contradiction arises and whether Hindi can in fact be the origin of this construction that seems to be so characteristic of Indian English. As this existential construction is attested almost exclusively in Indian English, it can be assumed that one or several of the background languages in India must be the source from which this feature enters the Indian variety of English. Future research will, however, be necessary to investigate which of the substrate languages are responsible for this word order in Indian existential constructions and exactly how this substrate accounts for the pattern in Indian English.

Let us now turn to Figure 27c. The corpus of Jamaican English shows an extremely strong preference of the collocation you have + indefinite noun phrase. It is also the only variety in which the collocation them have + indefinite noun phrase is attested at all. The prevalence of the first and the actual occurrence of the second can and must be attributed to the fact that, in Jamaican Creole, existence is normally expressed by constructions of the type you (h)av ‘you have’ or dem (h)av ‘they have’ (cf. Patrick 2007: 140). In informal Jamaican English, the creole pronoun dem is occasionally realized as them, even in subject position (see section 4.3 for the discussion of creole features in informal spoken Jamaican English) and, correspondingly, them have functions as a bridge pattern facilitating they-have-possessives in Jamaican English. These corpus findings thus provide evidence for a strong influence from the substrate language Jamaican Creole on Jamaican English with regard to possessive constructions used to express existence.

Last but not least, Singapore English is the variety with the highest frequency of got with an empty subject. According to Gupta (1994: 10f.), it is quite common on Colloquial Singapore English to omit the subject of the sentence when it is disambiguated by the context. This tendency toward PRO-drop utterances is particularly frequent in informal spoken language use and can be attributed to influence from Chinese. This is because Chinese is a null-subject language, i.e. a language which permits that an explicit subject
may be dropped when the topic of the sentence is otherwise made clear. In fact, “[i]n Chinese, both subjects and objects may drop from finite sentences” (Huang 1989: 186). Not only can the tendency to drop the subject be directly attributed to influence from Chinese, but so can the frequent use of got in Colloquial Singapore English for existential functions. Nomoto and Lee (2008) note that “got has a much wider range of meanings in Singlish than in standard varieties of English,” namely “possessive, existential, temporal location, aspect (habitual, experiential, completive), emphasis, challenge/disagreement (idiomatic), passive, ‘to receive/obtain’ and ‘to become’” (Nomoto and Lee 2008). In their 2007 article, Lee, Ling and Nomoto demonstrate that the main substrate accounting for the existential got construction in Colloquial Singapore English is southern varieties of Chinese, especially Hokkien (Lee, Ling and Nomoto 2007, see also Nomoto and Lee 2008). We can therefore conclude that the use of zero subject + got + indefinite noun phrase in Singapore English is another clear case of contact influence.

The next section is devoted to a discussion of potential explanations for variation in expressions of existence in the varieties under consideration, taking into account the findings presented in chapter 5 on agreement patterns in existential there + be constructions and the results of the investigation in chapter 6 of other existential constructions used in the respective Englishes.

6.4 Discussion – Universal Tendencies or Language Contact?

The comparative corpus investigation of agreement and non-agreement in existential there + be constructions in the native varieties British and American English and the postcolonial varieties Jamaican English, Indian English and Singapore English has revealed that register and the grammatical factor subject number exert a strong influence on agreement in all five varieties and thus constitute shared constraints applying across native and non-native varieties. By contrast, tense and polarity have been found to influence agreement in existential there + be constructions only in British English, but not in the L2 varieties.

The results have shown that it is rarely possible to draw clear distinctions between L2 varieties and L1 varieties only on the basis of a particular grammatical variable (or variables), in our case agreement patterns, as one of the L2s, Singapore English,

---

141 See Huang (1989) for a thorough investigation of PRO-drop phenomena in Chinese.
consistently patterns much more like the native varieties than like the other L2 varieties. The finding that Singapore English differs considerably from the other two non-native varieties would point to possible substrate influence, but on the other hand such language-contact influence is unlikely in view of the fact that usage patterns in the Singapore data consistently follow trends in the native Englishes, where no background languages are present. My results thus support the claim put forward by Sand that Singapore English might be “on the verge of becoming an L1-variety of English” (Sand 2005: 200). Similar observations are, of course, frequently made by observers of the sociolinguistic field in Singapore (e.g. Görlach (2002) and Foley (2001)), and it is interesting to see where precisely the sociolinguistic trend on the macro-level manifests itself on the micro-level of morphosyntactic structure. A thorough investigation of Singapore English in search of plausible explanations for the patterns found and an attempt to validate the above-mentioned assumption is unfortunately beyond the scope of the present work and will have to be postponed to future research.

Leaving aside Singapore English, a clear distinction emerges between variety types particularly for the overall amount of non-agreement and the grammatical-stylistic factor contractedness. The results confirm the sociolinguistic/ macro-sociological classification of Jamaican English as a post-colonial L2 English which resembles the L2 variety Indian English and evidently differs from both L1 varieties.

As most post-colonial contact varieties of English are not fully endonormative yet, it is of course interesting to see which of the globally dominant L1 standards – British or American – has the stronger influence. Historically, both Indian English and Jamaican English are uncontroversially British, but it is reasonable to speculate that, due to its geographical proximity and a longer and more intensive history of contact and migration, US English has played a greater part in shaping current norms of Jamaican usage. A recent study of modal-verb use in ICE-Jamaica (Mair 2009a), for example, has found Jamaican English to be relatively distant from British English. With regard to existential there + be constructions, however, Jamaican English patterns are consistently more similar to British than American English.

Jamaican English tends to follow the same trends as Indian English for all grammatical factors. As these L2s display such similar patterns, no strong case can be made for substrate influence and the more likely explanation is that angloversals determine the agreement patterns. Note, however, that the markedly language-specific nature of English existentials makes agreement in existential there + be constructions an unlikely point of
entry for direct contact influence. Although Jantos (forthc., and section 4.3.1) found subject-verb agreement with lexical verbs in non-existential sentences in Jamaican English to be strongly influenced by the creole substrate, with substrate-induced non-agreement (such as the frequent absence of third person –s on verbs) being particularly frequent in conversations, this trend is not reflected to a comparable extent in existential constructions in same text category. In fact, in terms of the frequency of non-agreement in existential there + be constructions there is no evidence at all of substrate influence in Jamaican English, nor in any other postcolonial variety. This shows that from the presence or absence of substrate influence in one construction, e.g. non-existentials, one cannot infer that such substrate influence is also present in other constructions, e.g. existential there + be constructions.

The investigation of non-canonical syntactic constructions used to express existence, however, reveals a substantial amount of evidence of substrate influence in the data of the four varieties analysed. The findings can be summarized as follows.

First, in the L2 Englishes there are hardly any instances of they’ve got or you’ve got + indefinite noun phrase, the non-canonical existential construction most frequently used in British English as an alternative to existential there + be constructions. This type of existential construction derives from possessive have got, which is an Early Modern English innovation and did not become frequent until the 18th century. This means that the further extension to existential constructions must have occurred even later (late 18th/19th centuries), and it is certainly interesting to see that such a relatively recent British innovation did not make it into either of the two historically British L2 varieties discussed here. In other words, with regard to existential-possessive constructions of the have got-type, the former colonial target variety British English does not influence the postcolonial varieties.

By contrast, strong preferences emerged in each post-colonial variety for other alternative expressions of existence, and these preferences were identified as clear cases of substrate influence.

In Jamaican English, the prevalence of you have + indefinite noun phrase can be traced back to a corresponding structure in the substrate, Jamaican Creole. That this is a case of contact influence is further supported by the fact that Jamaican English is the only variety in which the bridge pattern them have + noun phrase is attested at all. The use of them as

---

142 Due to the fact that the data basis of American English is not comparable to that of the other varieties, American English was excluded.
subject pronoun is derived from the creole indefinite subject pronoun *dem* (cf. Patrick 2007).

In the case of Indian English, evidence has been found for the presence of contact influence, but the source of this influence is not so clear. It seems that Hindi word order rules might be responsible for the high number of sentences in Indian English in which adverbiaal, sentence-final *there* is used instead of the common English dummy subject *there*. This is because Hindi normally distinguishes existential from non-existential sentences by means of word order, i.e. by “movement of the noun phrase from its canonical position” (Sinha and Thakur 2005: 249f.). However, the word order in Hindi existential constructions does not correspond to that found in the English data. While the fact that indefinite noun + *be* + *there* occurs almost exclusively in Indian English makes substrate sources likely, it is not clear how the application of word order movements in Hindi can account for the construction found so frequently in the Indian English data. Future research will have to shed more light on the substrate source of this particularly Indian existential construction.

In Singapore English, the tendency to drop subjects and the use of *got* for existential functions was significantly more frequent than in the two other varieties, and on the basis of previous research these features can be directly traced back to substrate influence from Chinese.

### 6.5 Conclusion

The investigation of existential constructions presented in chapters 5 and 6 provides strong evidence for the assumption that 1) vernacular universals, angloversals and substrate influence complement one another in accounting for the facts of existential constructions in native and non-native varieties of English, and 2) the relative strength of these competing factors is not constant across constructions but varies considerably.

Contact and substrate influence reveals itself to be weak to non-existent with regard to agreement in existential *there* + *be* constructions. This can be explained by the fact that existential constructions in the substrates in Jamaica, India and Singapore are structurally so different as to be practically irrelevant for (non-) agreement in existential *there* + *be* constructions in the respective English superstrates, and therefore the patterns are more likely caused by vernacular universals (e.g. variable agreement, register variation) and
angloversals (e.g. patterns relating to the grammatical factors subject number and contractedness).

On the other hand, substrate influence has been found to play a significant role in the frequencies with which non-canonical expressions of existence are used in the four varieties, such as locative structures, word order, and different types of possessive constructions.

In sum, the present investigation shows two things. First, the mere fact that English has subject-verb number agreement and a contact language (be it Jamaican Creole or Chinese) has not will not automatically lead to contact-induced language change in a particular construction, such as *there*-existentials. Second, however, the present paper has also provided evidence for the assumption that such contact influence is at work in non-canonical existentials, such as possessives of the *have*-type. This result is not unexpected because, after all, in bilingual situations it is not abstract decontextualised grammatical systems which are in contact, but concrete constructions used in authentic discourse by speakers and writers.
7 Concluding Remarks

The purpose of this study was to thoroughly investigate morphosyntactic variation in educated Jamaican English and to shed light on the complex net of factors that determine this variation. The linguistic variable selected for this study was subject-verb agreement, and the investigation was carried out from two perspectives.

In chapter 4, subject-verb agreement with lexical verbs and with the exceptional verb *be* was analysed across three spoken registers differing in their relative levels of formality and the results were compared to those obtained in a comparable study of three written registers of similar formality. The purpose of this detailed analysis of subject-verb agreement across six text categories in ICE-Jamaica was to determine the extent to which various linguistic and non-linguistic factors influence variety-internal variation. Among the linguistic factors discussed were, for example, the type of verb and the grammatical number of the subject, while the non-linguistic factors included register characteristics, such as the formality of the situation, and the mode of production, i.e. spoken or written. Of particular relevance were the following questions: 1) Are the different sociolinguistic conditions described in chapter 2, i.e. the creole continuum in the spoken domains versus diglossia in written language, indeed reflected in the pattern of subject-verb agreement? 2) How and to what extent are formality and the mode of production responsible for the agreement patterns in Jamaican English?

Answering these questions turned out to be a complicated matter, because the subject-verb agreement system in Jamaican English is very complex, with linguistic and non-linguistic factors strongly interacting in the creation of a multifaceted web of influences. The corpus analysis shows that formality and register characteristics seem to be better predictors of non-agreement than the mode of production, which has been found to exert only a limited influence on agreement patterns. With regard to formality, a stylistic continuum has been found in spoken and written language alike, with non-agreement rates being highest in informal language and decreasing with increasing formality. However, the spoken – written dichotomy also influences agreement patterns, albeit to a lesser extent. This is because, the results from the spoken data confirm the presence of the postulated creole continuum (cf. for example Deuber 2009a), with substrate influence being particularly strong and manifold in spoken informal Jamaican English and decreasing with formality. In written language, the amount of non-agreement is generally lower than in spoken...
language, and evidence of direct substrate influence is mainly restricted to informal texts. The situation can be better described as diglossia, with Standard English as the norm and only occasional appearances of overt creole features for rhetoric purposes (cf. Hinrichs 2006). What turned out to be of particular relevance, however, is the type of substrate influence. Overt creolisms, such as the use of mesoelectal pronouns (him, them as personal or possessive pronouns) or the extension of invariant copula (is/was) to plural contexts, are highly salient, infrequent (especially in written data) and decrease drastically with increasing formality. Less salient are zero forms, which are influenced by the general lack of inflection in the creole, e.g. uninflected third person singular verbs. These occur in all text categories across the whole data set, though with higher frequencies in spoken registers, and they tend to decrease with increasing formality. Even less salient and stigmatized are indirect influences from the creole which occur up into the registers of highest formality in both written and spoken language and go largely unnoticed even in written educational texts.

In chapter 5, the focus was shifted from intra-variety variation to a comparison of agreement patterns in one particular type of construction, existential there + be constructions, across five varieties of English. The characteristic patterns of (non-)agreement and their close relations with the grammatical factors subject number, tense, contractedness and polarity were analysed in the spoken section of four ICE-corpora and the SBC. The aim of this comparison was to explore differences and similarities in the agreement patterns of these varieties, in order to gain insights concerning the question how and to what extent substrate influence and universal tendencies contribute to the agreement in existential constructions.

The results show that the agreement patterns in this particular construction are not determined by substrate influence to any significant extent. Instead, three universal tendencies have been found. First, variation was found in all varieties under investigation. In other words, speakers of all Englishes have standard and non-standard variants at their disposal from which they select according to situation or grammatical circumstances. There is no variety of English in my sample in which either standard or non-standard agreement variants are applied categorically. This justifies the assumption that variable agreement may be a true universal in English.

Second, in all varieties considered, the analysis revealed a considerable amount of variation related to register. There is consistently more non-standard agreement in informal
text types than in highly formal language use. Register variation can thus also be considered as a candidate for true universals.

Third, the analysis indicated that agreement variation in existential *there* constructions is strongly determined by various grammatical constraints, such as contractedness or polarity, which apply differently to L1 and L2 Englishes. Although only five varieties were analysed in the present study, it is likely that these constraints are also at work in other varieties of English and might thus by universal tendencies applying differently to particular types of Englishes.

Jamaican English and Indian English were confirmed in their sociolinguistic status as post-colonial L2 Englishes, as they consistently behave differently from the L1 varieties British English and American English. The corpus study revealed surprising similarities in agreement patterns in existential *there* + *be* constructions between Jamaican and Indian English despite their different substrate languages. Substrate influence thus does not qualify as a plausible explanation and instead universal tendencies applying differently to the particular variety types must be responsible. The sociolinguistic status of Singapore English as a L2 variety, however, was not supported in my analysis as its patterns consistently resemble those of the L1s rather than the L2 varieties.

As there was hardly any trace of substrate influence in the particular syntactic phenomenon of agreement in existential *there* + *be* constructions, a further investigation of the ICE-corpora (i.e. of all varieties except American English) was carried out in chapter 6, in order to explore to what extent and in which ways the varieties apply alternative means of expressing existence, such as various types of possessive constructions and locative sentences. This corpus analysis revealed that each variety indeed has clear preferences for one or two particular constructions by means of which the existence of some entity can be expressed. These preferences can be accounted for by the respective background languages in India, Jamaica and Singapore and thus constitute clear cases of substrate influence. Again, there are direct and indirect contact influences.

In the continuum situation specific to Jamaica it is even possible to trace the stages of direct contact through what I have dubbed "bridge constructions," such as *them have*. This is the most common existential construction in Jamaican Creole. In informal Jamaican English in the ICE-data, *them have* is attested at modest frequencies but regularly, and it is very plausible to assume that this Jamaican Creole existential-possessive is the prime factor accounting for the higher frequency of its Jamaican English structural analogue *they have*. 

192
More frequent than such bridge constructions are zero forms, such as zero subject + got + indefinite noun phrase in Singapore English, which can be attributed to a tendency in Chinese to drop the subjects when they are disambiguated by the context. By far the most frequent type is indirect contact effects which manifest themselves in a contact-induced boost of the frequency of constructions which are already present in the target as marginal and non-canonical constructional types. An example of this is the high number of you have and they have-possessives in Jamaican English, whose use is facilitated by the presence of the creole constructions yu (h)av and dem (h)av, respectively.

Overall, the present study provides evidence for the fact that linguistic variation in Jamaica is caused by a multitude of factors. The characteristics of the register and mode of production undoubtedly play a very significant role in shaping the subject-verb agreement system in educated Jamaican English, but these alone do not explain the complex interaction of tendencies favouring or disfavouring standard agreement. As the comparative corpus study indicated, Mufwene’s (1986) postulation that universal tendencies and substrate influence complement one another in the creation of creole languages can be extended in its application to the corresponding acrolectal varieties, in our case the emerging standard of educated Jamaican English.

A very interesting and relevant large-scale investigation on the emerging standard variety of English in Jamaica was conducted by Deuber and has been completed only very recently. Deuber’s (2009b) point of departure is the observation in previous research (for instance Patrick 1999 and Winford 1972) that the English spoken in the Caribbean can be described as comprising “Creole-influenced varieties, but the exact nature of these varieties and how they relate to Standard English and the Creole itself are to be empirically investigated” (Deuber 2009b: 8). To do this is exactly the aim of her book “Style and standards in English in the Caribbean,” in which she compares morphological and syntactic variation, e.g. in direct wh-questions, copula forms, past marking of verbs, subject-verb agreement, verb negation and noun and pronoun morphology, in various spoken text categories of ICE-Jamaica with that in her corpus of Trinidadian English, as both are varieties of Caribbean English with a co-present creole language. Deuber is

---

143 She refers to her own investigation of the first 40 conversational texts in section S1A of ICE-Jamaica and uses, in addition, the results by Sand (1999) on radio and press texts for comparison when more formal text types are concerned.

144 Deuber is currently compiling an ICE-component for Trinidad and Tobago (ICE-T&T), and although this corpus has not yet reached completion she uses the available data on Trinidadian English as the basis of her comparison with ICE-Jamaica, namely 20 texts each from the categories of conversations, class lessons and broadcast news, and 15 unscripted speeches.
particularly interested in the special type of sociolinguistic variation that is characteristic of the Caribbean area especially with regard to style, but also standards.

Deuber (2009b) also compares her results from ICE-Jamaica and ICE-T&T with Patrick’s (1999) study of ten mesolectal speakers from the Kingston neighbourhood of Veeton and notes that the ICE-data cover only the upper range of the creole continuum, i.e. mainly acrolectal English, while Patrick’s Veeton data cover the whole mesolectal range. When the Veeton data are classified into a low, mid and high group respectively, Deuber’s conversational ICE-data resemble the Veeton high group in many respects, for instance with regard to past marking. Thus Deuber concludes that “[o]verall the ICE-Jamaica data are clearly distinct from the mesolectal Veeton data but can easily be linked to it via the upper mesolect” (2009b: 99). That is, there is a continuum of sociolinguistic variation that is based on the social status of the speakers, i.e. how high a degree of education they have reached. However, this continuum is also closely related to stylistic factors. Both, the present study and Deuber’s work, but also former research, among others Sand (1999), clearly prove that in conversational language in ICE-Jamaica more morphological and syntactic features of Jamaican Creole can be found than in more formal spoken text types such as radio broadcasts in case of Sand’s (1999) work or class lessons and broadcast news in case of the present investigation, and especially that such creole features are largely absent in written language of all formality levels, unless used consciously for specific rhetoric purposes (cf. also Hinrichs 2006). All these studies thus come to the conclusion that the notion of the creole continuum is a useful concept for the sociolinguistic variation found in spoken language, as it offers speakers “options for fine-tuning that go beyond anything that can be modelled against our current understanding of bilingualism, code-switching or code-mixing” (Mair 2002: 35), but it is not as useful a concept for descriptions of written language in Jamaica, where the diglossia model (cf. Mair 2002) has been found to be more adequate (see also Hinrichs 2006).

Deuber’s (2009b) study reveals a high amount of both inter- and intratextual variation in her data. Intertextual variation results in a “broad hierarchy of text categories in the spoken component of ICE-Jamaica in which levels of formality are correlated with the extent to which Creole features occur” (Deuber 2009b: 115), while intratextual variation means that acrolectal and mesolectal creole features can and do occur side by side within the same text, even in stretches of speech by the same speaker. On the basis of her results Deuber validates in principle Allsopp’s (1996) distinction between an “informal” level of usage in Caribbean English, “sometimes characterized by morphological and syntactic reductions of
English structure” (Allsopp 1996: lvi) and an “antiformal” level consisting of “any Creolized or Creole form or structure surviving or conveniently borrowed to suit context or situation” (Allsopp 1996: lvii), but she emphasizes that “informality in Jamaican English is a graded phenomenon, […] that the dividing line between what is informal and what is antiformal can be hard to draw (Deuber 2009b: 138).

Related to these usage levels is the “hierarchy among creole forms” proposed by Deuber (2009a: 29ff.). My thorough investigation of subject-verb agreement in various text categories of ICE-Jamaica has led me to extend this hierarchy and to postulate a ‘continuum of salience’ from highly salient and stigmatized overt creolisms at one end, such as the use of creole pronouns, e.g. him or them, as subject or possessive pronouns in English or the extension of invariant don’t or the invariant copula is/was to plural contexts, all of which occur mainly in informal spoken domains and are often used for specific rhetoric effects (and are thus found predominantly in Allsopp’s (1996) “antiformal level of usage”), over less salient zero forms, such as the absence of third person singular –s inflection, which occur in all spoken and written text types, and most frequently in informal registers, but which are stigmatized in registers of higher formality, to indirect contact influences, which are of so little salience as to go largely unnoticed even in texts of high formality, such as hypercorrectly inflected third person plural verbs.

Deuber adds that rare overt creole structures, such as a/naa as a preverbal aspect marker and the copula de “are examples of features which are mainly used anti-formally” (2009b: 140) in her data. She mentions wh-questions without inversion or do-support as a “prototypical example of an informal usage (2009b: 139), along with the absence of past marking (cf. Deuber 2009b: 99f.). Deuber’s and the present study both contribute to shedding light on the stylistic continuum characterising spoken educated Jamaican English.

An additional research question that Deuber considers very important and worth pursuing is “to what extent can morphological and syntactic variation in Caribbean English, especially non-standard usage in the more informal text categories, be explained by reference to the Creoles?” (Deuber 2009b: 18). She assumes that these creoles influence the sociolinguistic variation to a high extent, but Deuber emphasizes the importance of taking into account also potential angloversals (cf. Chambers 2004, Mair 2003, Sand 2005) or effects of second-language learning typical in ESL contexts (cf. Sand 2005) in explanations of such synchronic variation.

Deuber’s analysis of the morphology and syntax of Jamaican English and the comparison with results from her Trinidadian English data also shows that substrate influence and
angloversals both contribute to the complex pattern of sociolinguistic variation in Caribbean English, a result that is very similar to that of the investigation described in the present work.

It was noted above in the comparative study of existential *there + be* constructions that certain agreement patterns seem to be shared among the L2 varieties analysed, i.e. Singapore English, Indian English and Jamaican English. For example, there is variable agreement in existential *there* constructions in all the varieties studied which is based on register characteristics and formality as well as on grammatical factors such as subject number.

Similarly, Platt, Weber and Ho (1984), Sand (2005), Kortmann and Szmrecsanyi (2004) and Deuber (2009b) find that some features seem to be shared by many New Englishes, for instance, the ESL varieties represented by ICE-components. Sand lists omission of articles, lacking inflection in plural nouns or past tense verbs as “transparent features, which are also quite salient especially if used in writing” (2005: 198) and gives noun reclassifications, blurred distinctions between past tense, present and past perfect and progressives with stative verbs, among others, as “opaque features (cf. Shastri 1992: 269f.), which are only recognised in comparison to data from other varieties of English because they are not differences in form, but rather in usage conventions” (Sand 2005: 198). A comparison of Jamaican data with ICE-India reveals that *wh*-questions lacking inversion or *do*-support are common in both samples and thus indeed seem to be a characteristic feature of ESL varieties (cf. Deuber 2009b). On the other hand, Deuber finds that features such as zero copula or unmarked third person singular verbs are not at all characteristic of an ESL variety like educated Indian English in the way they are of informal Jamaican English as represented in the present conversation sample [i.e. ICE-Jamaica, S1A-001-040, S. J.]. This can be explained by the fact that these forms have a different status in both contexts, i.e. what are basically learner errors in Indian English are quite appropriate forms in some contexts in Jamaican conversations because of the way English and Creole are integrated at this level of spoken language use. (Deuber 2009b: 141f.)

Some features that the present author assumes to be a result of influence from the creole in Jamaica have also been analysed comparatively by Deuber (2009b) in Jamaican and Trinidadian English. One of them is the tendency in Jamaican English to use *don’t* in third person singular contexts which has been attributed to the mesolectal Jamaican invariant negator *duon*. Deuber’s Trinidadian data reveals no such tendency, and this is most likely
so because “Trinidadian Creole has no general negator equivalent to duon. Instead it has several forms for main verb negation in the present tense, namely […] ain’t, don’t and doesn’t (both of the latter insensitive to person/number), so generalization of don’t is less likely” (Deuber 2009b: 162).

Another of Deuber’s findings is that zero copula are much more frequent in all locations in Trinidadian as compared to Jamaican conversations. While the analysis does not fully explain the differences in frequency rates of zero copula between the two conversational data sets, Deuber notes that “the findings from the ICE-T&T data certainly strongly support the main conclusion […] that zero copula use in acrolectal speech has strong roots in the Creole” (2009b: 153) in Trinidadian English as well as Jamaican English.

The analysis of non-canonical existential constructions (cf. chapter 6), such as the Jamaican English bridge construction them have facilitating ‘have-existentials’, the apparent tendency in Indian English to express existence by means of word order, and the ‘got-existentials’ prevalent in Singapore English, again make a strong point for the influence exerted in one way or another by contact with other languages used in the respective region. At the same time the results indicate that rates of agreement and non-agreement in existential there + be constructions are not influenced in any of the varieties considered by the presence or absence of particular types of substrate languages, in fact the crucial determining factors turned out to be grammatical ones, such as subject number.

With regard to substrate influence Sand notes that “the presence or absence of a feature in the relevant substrate(s) is only of limited influence in the contact varieties. It may best be described as a possible reinforcement of tendencies already at work and its influence is most pronounced in informal conversations” (2005: 201). This statement seems to contradict Deuber’s (2009b) conclusion, but the data sets and the morphological and syntactic variables these linguists analysed differed considerably, as Deuber explicitly notes: “the data and variables chosen in order to investigate the central topic of style quite obviously bias the view towards those aspects of morphological and syntactic variation in English in the Caribbean that are most susceptible to the influence of the creoles”.

All in all, therefore, the attempt to explain variation in L2 varieties by reference to angloversals and language contact is by no means an ‘either—or question’, it is clear that both angloversals AND language contact have strong effects on linguistic variation. What still needs to be explained in future research by means of thorough investigation is the way in which they interact and the extent to which each of them plays a role in specific
linguistic variables and at particular levels of formality, a conclusion on which both Deuber (cf. 2009b: 258) and the author of the present study agree.

Last, Deuber addresses the question of the sociolinguistic status of Jamaican English. She points to a study by Mair (2009b) investigating be-contractions in the conversational sections of the ICE-components from Great Britain, New Zealand, Jamaica and India. Mair shows that such be-contractions tend to indicate an informal style in ENL varieties. His results indicate that Jamaican English contraction rates are lower than those found in the native speaker corpora, but they remain “nevertheless much closer to them than to a clear second-language variety such as Indian English” (Mair 2009: 21). That means the Jamaican rates resemble those of ENL varieties such as British English and New Zealand English much more than those of ESL varieties such as Indian English.

This is in stark contrast to the results obtained in the investigation of existential there + be constructions (see chapter 5), where Jamaican English consistently patterns more like L2 varieties than L1. Taking into account the finding that features are attested in Caribbean varieties which are common in pidgin and creole languages (cf. Kortmann and Szmrecsanyi 2004: 1191), less frequent in ESL varieties and quite rare in ENL countries, Deuber emphasises that “English in the Caribbean comprises informal varieties that have a highly special character, sharing features with ESL varieties, ENL varieties, and Pidgins/Creoles, without really belonging in either of these groups” (Deuber 2009b: 259). Hence she concludes that “the use of ESD as a specific label for such varieties seems highly justified” (p. 259).

Not only the status of Jamaican English as ESL variety can be doubted on the basis of recent research, Singapore English too seems to be a less clear case of ESL than previously assumed, as discussed above with regard to agreement in existential there + be constructions. Sand also questions the status of Singapore English as ESL, she assumes that Singapore English might be “on the verge of becoming an L1-variety of English” (2005: 200; see also Leimgruber’s (2008) insightful study Sociolinguistic Variation in Singapore English).

Most of the studies mentioned here, including the present investigation, have profited greatly from the availability of comparative data sets for a considerable number of varieties of English in the ICE-project.

Using national components of the ICE family as the basis for the present study showed that the corpus project is extremely valuable to the linguist conducting comparative investigations of several varieties of English. The main advantages of the corpus project
are 1) that it provides parallel data sets for a large number of varieties and therefore ensures the comparability of research results across these corpora, 2) that it provides data for L1 varieties such as British English, Australian English or New Zealand English as well as for L2 varieties such as Jamaican English, Singapore English or Indian English and thus has made the present cross-variety investigation possible in the first place, 3) that both spoken and written text types are included in a ratio of 300 : 200 texts per corpus, so that the mode of production on linguistic variables can be studied and 4) that there are also a number of different text categories which cover a large range on the stylistic continuum, from quite informal conversational texts to highly formal press and news texts. The text categories in the national component corpora all comprise equal numbers and sizes of texts, and, wherever possible, a similar thematic make-up. This strict classification of texts into categories is very useful when it comes to investigations of stylistic variation or the dependence of variables on register characteristics. However, there are also some downsides to the corpora. First, some of the text categories are rather heterogeneous, so that it is sometimes difficult to come to reliable conclusions about the effects of register characteristics on linguistic variables such as subject-verb agreement and non-agreement. The spoken category ‘class lessons’, for example, consists of very lively and rather informal discussions between students and their teachers as well as much more formal lectures, and the written category ‘social letters’ comprises personal email communication as well as more formal business letters. For this reason, the individual categories are sometimes hard to distinguish from one another with regard to particular features, for instance formality.

The second disadvantage from the point of view of the present study is the size of the corpora. The samples of one million words per corpus are large enough for studies of highly frequent phenomena such as subject-verb agreement and non-agreement in lexical verbs, but many interesting questions cannot be answered on the basis of these corpora, as there are simply not enough tokens in the data. In some cases the numbers suffice to make tentative statements about possible correlations but are too low for reliable statistical significance tests, for instance in case of the different types of non-agreement with lexical verbs and the influence of register characteristics on their frequency patterns in different spoken and written text categories.

Studying the influence of polarity on agreement with the verb be, or conducting a comparative analysis of agreement variation in existential there + be constructions across various spoken and written text categories and its dependence on grammatical factors such
as contractedness, polarity, and particularly the influence of particular negative determiners on non-agreement in sentences of negative polarity, are only few of the endeavours which proved to be impossible due to insufficient token numbers in the data and thus had to be given up.

Third, some interesting phenomena are hard to search by means of the WordSmith program as the ICE-corpora, with the single exception of ICE-GB, are not tagged for parts of speech. Unfortunately, it was necessary to manually mark all lexical verbs in present tense for agreement or non-agreement, so that only a restricted number of text categories could be chosen for the study and only a rather low number of texts per category could be included, e.g. only 20 out of 100 conversational texts were analysed.

Another disadvantage is that only studies of educated English are possible in the ICE-corpora, due to the strict criteria for speaker inclusion. Only adult speakers who have completed secondary education are included, unless their social position as politician or other important member of the public warrants their inclusion. This means that only a restricted range of the social-stylistic continuum can be investigated. In the case of Jamaica this means that only the upper, i.e. acrolectal and upper mesolectal, range of the creole continuum between educated Jamaican English at the one end and Jamaican Creole at the other can be investigated and other sources have to be used for studies of the characteristics of lower mesolectal and basilectal language usage.

Last, the fact that L2 varieties tend to use different codes for different functions, namely English for formal functions such as administration, commerce and education, and one of the local languages for the private sphere and informal communication, for instance the creole in Jamaica, Hindi in India, and Mandarin Chinese in Singapore, results in a bias towards formal language use in the L2 corpora, as the use of English automatically implies a higher level of formality than for example in British English. This can be a problem for studies of L1 and L2 varieties as it endangers the comparability of the results.

What the present work also shows is that the Santa Barbara Corpus of Spoken American English is not a suitable basis for comparative analyses with ICE-corpora. This is because it was compiled differently from the other corpora. It presently consists of only 83 texts from spoken informal conversations and phone calls, as compared to 300 spoken texts in the other ICE-corpora, 100 of which are conversational texts, but which also include 200 texts from a number of other spoken text types such as class lessons and broadcast news. Despite these disadvantages, the ICE project is an immensely useful and valuable resource for the linguist interested in comparative investigations of different varieties of English.
Subject-verb agreement, the variable analysed in this work, is, of course, only one of the countless pieces in the puzzle of the morphosyntax of Jamaican English, and more research of the kind provided by Sand (1999 and 2005), Deuber (2009a and b) or Mair (2002, 2003) or the present work is still needed in order to gain deeper insights into the interaction of the many factors contributing to its characteristic appearance. Once a sub-corpus of American English is released as part of the ICE project, a follow-up study would be valuable, in order gain a clearer picture on potential influences of American English on Jamaican English, and a similar investigation as the one presented here could be carried out for further morphosyntactic phenomena. Moreover, scholars are encouraged to compare a larger number of L1 and L2 Englishes in order to further validate the claims put forward in this book on the sociolinguistic status of Jamaican English as L2 variety and that of Singapore English as L1 variety ‘in the making.’

In this way, research on the ICE family of corpora will be able to systematically expand our understanding of the complex interplay of universal or at least pan-English factors and contact in the shaping of morphosyntactic variation.
Appendix

1. Text Categories in ICE-Jamaica

<table>
<thead>
<tr>
<th>Spoken (300)</th>
<th>Dialogues (180)</th>
<th>Private (100)</th>
<th>Conversations (90)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Public (80)</td>
<td></td>
<td>Phonecalls (10)</td>
</tr>
<tr>
<td></td>
<td>Monologues (120)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unscripted (70)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Scripted (50)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Written (200)</td>
<td>Non-printed (50)</td>
<td>Student Writing (20)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Letters (30)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Printed (150)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Academic (40)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Popular (40)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reportage (20)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Instructional (20)</td>
<td></td>
<td>Administrative Writing (10)</td>
</tr>
<tr>
<td></td>
<td>Persuasive (10)</td>
<td></td>
<td>Skills/hobbies (10)</td>
</tr>
<tr>
<td></td>
<td>Creative (20)</td>
<td></td>
<td>Novels (20)</td>
</tr>
</tbody>
</table>

Class Lessons (20)  
Broadcast Discussions (20)  
Broadcast Interviews (10)  
Parliamentary Debates (10)  
Cross-examinations (10)  
Business Transactions (10)  
Commentaries (20)  
Unscripted Speeches (30)  
Legal Presentations (10)  
Broadcast News (20)  
Broadcast Talks (20)  
Non-broadcast Talks (10)  
Student Essays (10)  
Exam Scripts (10)  
Social Letters (15)  
Business Letters (15)  
Humanities (10)  
Social Sciences (10)  
Natural Sciences (10)  
Technology (10)  
Humanities (10)  
Social Sciences (10)  
Natural Sciences (10)  
Technology (10)  
Press reports (20)  
Administrative Writing (10)  
Editorials (10)  
Novels (20)
# 2. Demographic Characteristics and Linguistic Make-Up of the Population of Singapore 1990 and 2000

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Singapore Residents ('000)</td>
<td>2,735.9</td>
<td>3,263.2</td>
<td>2,127.9</td>
<td>2,505.4</td>
<td>384.3</td>
<td>453.6</td>
<td>194.0</td>
<td>257.8</td>
<td>29.6</td>
<td>46.4</td>
</tr>
<tr>
<td>Ethnic Composition (%)</td>
<td>100.0</td>
<td>100.0</td>
<td>77.8</td>
<td>76.8</td>
<td>14.0</td>
<td>13.9</td>
<td>7.1</td>
<td>7.9</td>
<td>1.1</td>
<td>1.4</td>
</tr>
<tr>
<td>Median Age (Years)</td>
<td>29</td>
<td>34</td>
<td>30</td>
<td>35</td>
<td>26</td>
<td>29</td>
<td>29</td>
<td>33</td>
<td>29</td>
<td>33</td>
</tr>
<tr>
<td>Age Dependency Ratio (Per 100 Adults 15 – 64 Years)</td>
<td>40.8</td>
<td>40.4</td>
<td>39.2</td>
<td>37.4</td>
<td>49.4</td>
<td>54.3</td>
<td>40.8</td>
<td>46.6</td>
<td>54.0</td>
<td>48.9</td>
</tr>
<tr>
<td>Under 15</td>
<td>32.3</td>
<td>30.1</td>
<td>30.4</td>
<td>27.0</td>
<td>43.0</td>
<td>45.6</td>
<td>32.9</td>
<td>35.4</td>
<td>42.3</td>
<td>37.7</td>
</tr>
<tr>
<td>65 &amp; Over</td>
<td>8.5</td>
<td>10.2</td>
<td>8.8</td>
<td>10.4</td>
<td>6.4</td>
<td>8.8</td>
<td>7.9</td>
<td>11.2</td>
<td>11.7</td>
<td>11.2</td>
</tr>
<tr>
<td>Sex Ratio (Males Per 1,000 Females)</td>
<td>1,027</td>
<td>998</td>
<td>1,013</td>
<td>989</td>
<td>1,040</td>
<td>1,012</td>
<td>1,179</td>
<td>1,092</td>
<td>933</td>
<td>885</td>
</tr>
<tr>
<td>Literacy &amp; Language (%)</td>
<td>89.1</td>
<td>92.5</td>
<td>88.5</td>
<td>92.1</td>
<td>90.3</td>
<td>93.6</td>
<td>92.0</td>
<td>95.1</td>
<td>96.3</td>
<td>97.1</td>
</tr>
<tr>
<td>General Literacy Rate (Aged 15 Years &amp; Over)</td>
<td>89.1</td>
<td>92.5</td>
<td>88.5</td>
<td>92.1</td>
<td>90.3</td>
<td>93.6</td>
<td>92.0</td>
<td>95.1</td>
<td>96.3</td>
<td>97.1</td>
</tr>
<tr>
<td>Language Most Frequently Spoken At Home (Aged 5 Years &amp; Over)</td>
<td>English</td>
<td>18.8</td>
<td>23.0</td>
<td>19.3</td>
<td>23.9</td>
<td>6.1</td>
<td>7.9</td>
<td>32.3</td>
<td>35.6</td>
<td>67.2</td>
</tr>
<tr>
<td>Mandarin</td>
<td>23.7</td>
<td>35.0</td>
<td>30.1</td>
<td>45.1</td>
<td>-</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>2.8</td>
<td>4.4</td>
</tr>
<tr>
<td>Chinese Dialects</td>
<td>39.6</td>
<td>23.8</td>
<td>50.3</td>
<td>30.7</td>
<td>-</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>3.3</td>
<td>3.2</td>
</tr>
<tr>
<td>Malay</td>
<td>14.3</td>
<td>14.1</td>
<td>0.3</td>
<td>0.2</td>
<td>93.7</td>
<td>91.6</td>
<td>14.5</td>
<td>11.6</td>
<td>19.3</td>
<td>15.6</td>
</tr>
<tr>
<td>Tamil</td>
<td>2.9</td>
<td>3.2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>Others</td>
<td>0.8</td>
<td>0.9</td>
<td>-</td>
<td>0.1</td>
<td>-</td>
<td>0.3</td>
<td>9.8</td>
<td>9.7</td>
<td>7.3</td>
<td>8.2</td>
</tr>
</tbody>
</table>

### 3. Scheduled Languages in India in Descending Order of Speakers’ Strength – 2001

<table>
<thead>
<tr>
<th>No.</th>
<th>Language</th>
<th>Persons who returned the language as their mother tongue</th>
<th>Percentage of total population*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hindi</td>
<td>422,048,642</td>
<td>41.03</td>
</tr>
<tr>
<td>2</td>
<td>Bengali</td>
<td>83,369,769</td>
<td>8.11</td>
</tr>
<tr>
<td>3</td>
<td>Telugu</td>
<td>74,002,856</td>
<td>7.19</td>
</tr>
<tr>
<td>4</td>
<td>Marathi</td>
<td>71,936,894</td>
<td>6.99</td>
</tr>
<tr>
<td>5</td>
<td>Tamil</td>
<td>60,793,814</td>
<td>5.91</td>
</tr>
<tr>
<td>6</td>
<td>Urdu</td>
<td>51,536,111</td>
<td>5.01</td>
</tr>
<tr>
<td>7</td>
<td>Gujarati</td>
<td>46,091,617</td>
<td>4.48</td>
</tr>
<tr>
<td>8</td>
<td>Kannada</td>
<td>37,924,011</td>
<td>3.69</td>
</tr>
<tr>
<td>9</td>
<td>Malayalam</td>
<td>33,066,392</td>
<td>3.21</td>
</tr>
<tr>
<td>10</td>
<td>Oriya</td>
<td>33,017,446</td>
<td>3.21</td>
</tr>
<tr>
<td>11</td>
<td>Punjabi</td>
<td>29,102,477</td>
<td>2.83</td>
</tr>
<tr>
<td>12</td>
<td>Assamese</td>
<td>13,168,484</td>
<td>1.28</td>
</tr>
<tr>
<td>13</td>
<td>Maithili</td>
<td>12,179,122</td>
<td>1.18</td>
</tr>
<tr>
<td>14</td>
<td>Santali</td>
<td>6,469,600</td>
<td>0.63</td>
</tr>
<tr>
<td>15</td>
<td>Kashmiri</td>
<td>5,527,698</td>
<td>0.54</td>
</tr>
<tr>
<td>16</td>
<td>Nepali</td>
<td>2,871,749</td>
<td>0.28</td>
</tr>
<tr>
<td>17</td>
<td>Sindhi</td>
<td>2,535,485</td>
<td>0.25</td>
</tr>
<tr>
<td>18</td>
<td>Konkani</td>
<td>2,489,015</td>
<td>0.24</td>
</tr>
<tr>
<td>19</td>
<td>Dogri</td>
<td>2,282,589</td>
<td>0.22</td>
</tr>
<tr>
<td>20</td>
<td>Manipuri**</td>
<td>1,466,705</td>
<td>0.14</td>
</tr>
<tr>
<td>21</td>
<td>Bodo</td>
<td>1,350,478</td>
<td>0.13</td>
</tr>
<tr>
<td>22</td>
<td>Sanskrit</td>
<td>14,135</td>
<td>N</td>
</tr>
</tbody>
</table>

*The percentage of speakers of each language for 2001 has been worked out on the total population of India excluding the population of Mao-Maram, Paomata and Purul subdivisions of Senapati district of Manipur due to cancellation of census results. N - Stands for negligible.

** Excludes figures of Paomata, Mao-Maram and Purul sub-divisions of Senapati district of Manipur for 2001.

References


Hinrichs, Lars. 2006. *Codeswitching on the Web: English and Jamaican Creole in E-mail Communication.* (Pragmatics and Beyond 147). Amsterdam/Phil.: Benjamins.


Schneider, Edgar W. 2000. “From region to class to identity: ‘Show me how you speak, and I’ll tell you who you are?’” *American Speech* 75, 359-361.


Zusammenfassung in deutscher Sprache


Im Zentrum der Untersuchung stehen Umfang und Intensität des Einflusses, den das JamC auf das JamE im Sprachkontakt ausübt, und zwar in Abhängigkeit von Formalität der Situation und Produktionsmodalität (gesprochen oder geschrieben). Es stellt sich einerseits
heraus, dass die Formalität insgesamt einen stärkeren Einfluss auf das Kongruenzsystem hat als die Produktionsmodalität, und andererseits, dass direkte Einflüsse aus dem Substrat, JamC, sich fast ausschließlich auf informelle gesprochene Sprache beschränken. Der nach Formalitätsgrad variable Substrateinfluss führt zu einem hohen Maß an Komplexität im jamaikanischen Kongruenzsystem, das in der Arbeit detailliert beschrieben wird.

Für einen Teil des untersuchten Phänomenbereichs – nämlich Kongruenz bei Existentialkonstruktionen des Typs \( \text{there} + \text{be} \) – liefert die Arbeit auch varietätenvergleichende Analysen auf der Grundlage weiterer ICE-Corpora (GB, Singapore, India) sowie eines amerikanischen Vergleichskorpus. Diese komparative Untersuchung erfolgt in Kapitel 5, das mit einer kurzen Zusammenfassung bisheriger Forschungsergebnisse zu \( \text{there} + \text{be} \) Konstruktionen beginnt, anschließend die zu untersuchenden Varietäten, ihre jeweilige Entstehungsgeschichte und ihre gegenwärtige soziolinguistische Situation vorstellt, die Methodik kurz beschreibt, und schließlich zur vergleichenden Untersuchung der Kongruenzmuster bei dieser Existentialkonstruktion übergeht. Hierbei zeigt sich, dass die Häufigkeit von (In)kongruenz in hohem Maße an grammatische Faktoren gebunden ist und sich Unterschiede eher auf den Varietätentyp (L1 versus L2) als auf Substrateinfluss zurückführen lassen.

In Kapitel 6 wird der Fokus in zweierlei Hinsicht erweitert. Erstens werden in einer typologischen Diskussion Strukturen thematisiert, die als Alternativen zu \( \text{there} + \text{be} \) Konstruktionen zum Ausdrücken von Existenz dienen können, wie zum Beispiel Lokativ- und Possessivkonstruktionen (6.1), und zweitens wird der Analysefokus auf die jeweiligen Substrate in Jamaika, Indien und Singapur ausgeweitet, um die dort gebräuchlichen Ausdrücke zu identifizieren (6.2). Im Anschluss folgt eine Korpusanalyse der entsprechenden Existentialkonstruktionen in vier ICE-Korpora (das amerikanische Corpus wird auf Grund mangelnder Vergleichbarkeit nicht einbezogen), um möglichen Substrateinfluss auf die jeweiligen Varietäten aufzudecken (6.3). Die Korpusanalyse offenbart starke Präferenzen in den einzelnen Varietäten für unterschiedliche Konstruktionen, die sich eindeutig Kontakteinflüssen aus den Substratsprachen zuordnen lassen. Was die Kontaktvarietäten (JamEng, Indian English, Singapore English) betrifft, zeigt sich der Einfluss der jeweiligen Kontaktsprachen eher in der Präferenz für andere Formen der Existentialkonstruktion (z.B. mit \( \text{have} \)) als in der Häufigkeit von (In)kongruenz bei \( \text{there} + \text{be} \).

In einer abschließenden Diskussion (7.) werden die wichtigsten Ergebnisse noch einmal zusammengefasst und in den Kontext aktueller Forschung gesetzt.